

ALESSANDRA POPOV DOS SANTOS MANHA

**Análise da expressão diferencial em
três fases da esporulação de
oocistos de *Eimeria maxima***

Tese apresentada ao Programa de Pós-Graduação em Biologia da Relação Patógeno-Hospedeiro do Instituto de Ciências Biomédicas da Universidade de São Paulo, para obtenção do Título de Doutor em Ciências.

**São Paulo
2011**

ALESSANDRA POPOV DOS SANTOS MANHA

**Análise da expressão diferencial em
três fases da esporulação de
oocistos de *Eimeria maxima***

Tese apresentada ao Programa de Pós-Graduação em Biologia da Relação Patógeno-Hospedeiro do Instituto de Ciências Biomédicas da Universidade de São Paulo, para obtenção do Título de Doutor em Ciências.

Área de Concentração: Biologia da Relação Patógeno-Hospedeiro.

Orientador:
Prof. Dr. Alda Maria Backx Noronha Madeira

**São Paulo
2011**

DADOS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
Serviço de Biblioteca e Informação Biomédica do
Instituto de Ciências Biomédicas da Universidade de São Paulo

© reprodução total

Manha, Alessandra Popov dos Santos.

Análise da expressão diferencial em três fases da esporulação de oocistos de *Eimeria maxima* / Alessandra Popov dos Santos Manha. -- São Paulo, 2011.

Orientador: Alda Maria Backx Noronha Madeira.

Tese (Doutorado) – Universidade de São Paulo. Instituto de Ciências Biomédicas. Departamento de Parasitologia. Área de concentração: Biologia da Relação Patógeno-Hospedeiro. Linha de pesquisa: Biologia Molecular de *Eimeria*.

Versão do título para o inglês: Differential expression analysis of *Eimeria maxima* oocysts in three phases of sporulation.

Descritores: 1. *Eimeria maxima* 2. Coccidiose 3. Oocisto 4. Esporulação 5. Expressão gênica 6. ORESTES I. Madeira, Alda Maria Backx Noronha II. Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação em Parasitologia III. Título.

ICB/SBIB025/2011

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS

Candidato(a): Alessandra Popov dos Santos Manha.

Título da Tese:
esporulação Análise da expressão diferencial em três fases da
de oocistos de *Eimeria maxima*.

Orientador(a): Madeira, Alda Maria Backx Noronha.

A Comissão Julgadora dos trabalhos de Defesa da Tese de Doutorado, em sessão pública realizada a/...../....., considerou

() Aprovado(a)

() Reprovado(a)

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Presidente: Assinatura:
Nome:
Instituição:




**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS**

Cidade Universitária "Armando de Salles Oliveira"
Av. Prof. Lineu Prestes, 2415 – CEP. 05508-000 São Paulo, SP – Brasil
Telefone : (55) (011) 3091.7733 – telefax : (55) (011) 3091.7438
e-mail: cep@icb.usp.br


CERTIFICADO

Certificamos que o protocolo registrado sob nº **124** nas fls. **10** do livro **2** para uso de animais em experimentação, sob a responsabilidade da Profa. Dra. Alda Maria Backx Noronha Madeira, Coordenadora da Linha de Pesquisa "**Biologia Molecular e celular de parasitas de galinha doméstica do gênero etmeria**" do qual participou(aram) o(s) aluno(s): **Jeniffer Novaes Gonçalves Dias, Alessandra Popov dos Santos Manha** e o pesquisador **Arthur Gruber**, está de acordo com os Princípios Éticos de Experimentação Animal adotado pelo Colégio Brasileiro de Experimentação Animal (COBEA) e foi aprovado pela **COMISSÃO DE ÉTICA EM EXPERIMENTAÇÃO ANIMAL (CEEA)** em **17.02.2005**.

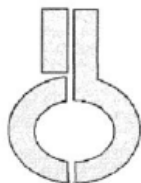
São Paulo, 18 de fevereiro de 2005.



Profa. Dra. Marília C. Leite Seelaender
Coordenadora da CEEA



Prof. Dr. Francisco Carlos Pereira
Secretário da CEEA



**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS**

Cidade Universitária "Armando de Salles Oliveira"
Av. Prof. Lineu Prestes, 2415 - cep. 05508-000 São Paulo, SP - Brasil
Telefone : (55) (011) 3091.7733 - telefax : (55) (011) 3091 7438
e-mail: cep@icb.usp.br

Of.CEEA.017.08
WTL/mcgn

São Paulo, 11 de novembro de 2008.

REF.: Protocolo nº 124/04.

"Biologia Molecular e celular de parasitas de galinha doméstica do gênero Eimeria"

Prezada Professora,

Informo que a sua licença para uso de animais em experimentação, constante no protocolo em epígrafe, foi prorrogada até 17.02.2011.

Reitero que havendo alteração de metodologia e inserção de novos alunos ao projeto de pesquisa vinculado à referida licença a CEEA/ICB deverá ser informada.

Cordialmente,

Prof. Dr. WOTHAN TAVARES DE LIMA
Coordenador da Comissão de Ética em
Experimentação Animal - ICB /USP

Ilma. Sra.

Profa. Dra. ALDA MARIA BACKX NORONHA MADEIRA
Departamento de Parasitologia
Instituto de Ciências Biomédicas -USP

Dedico esta tese aos meus pais, meu marido e ao meu irmão pelo apoio incondicional, pelo amor e respeito, presente em todos os momentos.

AGRADECIMENTOS

Ao finalizar esta etapa particularmente importante da minha vida, não poderia deixar de expressar o mais profundo agradecimento a todos que me apoiaram nesta longa caminhada e contribuíram para a realização deste trabalho.

Agradeço primeiramente a minha querida orientadora, Alda, que sempre demonstrou acreditar em meu potencial, pela oportunidade oferecida, pela estrutura do laboratório, pela orientação e principalmente pelo bom convívio nestes anos de trabalho. Com ela tive a oportunidade de enriquecer meu conhecimento, com suas argumentações científicas e sugestões durante o projeto.

Ao professor Arthur por auxiliar nas questões científicas e contribuir com seu conhecimento em análises durante todo este projeto.

Ao professor Alan pelo auxílio no desenvolvimento dos programas de bioinformática utilizados neste projeto.

Ao professor Paulo por me auxiliar nas análises do programa Threader.

A Jeniffer que acreditou em meu potencial desde a iniciação científica, me levando a fazer este doutorado. Além de contribuir com meu enriquecimento intelectual, me auxiliando em muitas análises realizadas neste projeto. Agradeço também pelo companheirismo e amizade durante todos estes anos.

A Ursa pela amizade e companheirismo no laboratório e fora dele, auxiliando em todos os momentos difíceis.

A Milene pela criação do protocolo de anotação automática e pela geração destes dados, utilizados neste projeto.

A todos os membros do laboratório, Laure, Thibério, André, Marcelo, Érika, além de todas as outras pessoas que já passaram por este laboratório, contribuindo com momentos alegres no ambiente de trabalho.

A minha primeira orientadora, Isabela que permitiu minha familiarização com laboratório de pesquisa, além de ensinar toda metodologia científica necessária durante um projeto.

A Judite por me iniciar na carreira científica, e pelo companheirismo.

Aos meus pais, pelos esforços e incentivos durante toda minha vida, além do amor incondicional.

Ao meu irmão por me ajudar nas horas em que precisei.

A minha cunhada Alessandra pelos momentos alegres e de incentivo.

Ao meu marido, Marcelo, pelos momentos de compreensão, amor, carinho e por sempre estar disposto a me ajudar em qualquer situação.

A toda minha família pelo carinho e incentivo.

A todos os meus amigos, principalmente as minhas melhores amigas Julianas (Oneda e Makita) e a Vanessa que fizeram parte de todos os momentos da minha vida me ajudando e incentivando.

À Granja Kunitomo, pela doação das aves utilizadas na propagação das espécies de *Eimeria* de galinha.

A cada componente desta banca, por participar do começo de mais uma etapa na minha vida.

Ao Instituto de Ciências Biomédicas da Universidade de São Paulo, pela formação durante a pós-graduação.

A CAPES e Conselho Nacional de Pesquisa (CNPq) pelas bolsas de estudos concedidas).

A Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) pelo apoio financeiro ao laboratório e ao presente projeto.

“Para mim, é muito melhor compreender o universo como ele realmente é, do que persistir no engano, por mais satisfatório e tranquilizador que possa parecer.”

Carl Sagan

RESUMO

Manha APS. Análise da expressão diferencial em três fases da esporulação de oocistos de *Eimeria maxima*. [tese (Doutorado em Biologia da Relação Patógeno-Hospedeiro)]. São Paulo: Instituto de Ciências Biomédicas da Universidade de São Paulo; 2011.

A coccidiose aviária é causada por protozoários do gênero *Eimeria*, sendo a *E. maxima* uma das espécies mais patogênicas e de larga distribuição mundial. Visando uma maior compreensão dos mecanismos moleculares envolvidos na esporulação em *E. maxima*, o perfil da expressão gênica foi avaliado em oocistos não esporulados, parcialmente esporulados e esporulados empregando a técnica ORESTES (*Open Reading Frame Expressed Sequence Tags*). Esta metodologia possibilita obter sequências transcritas parciais derivadas majoritariamente da porção central dos RNAs mensageiros, favorecendo a descoberta de novos genes. Foram geradas 20.149 leituras totais e após o pré-processamento, 75% das sequências foram aceitas. Posteriormente ao agrupamento das sequências, foram obtidos 1.207 *contigs* e 2.091 *singlets*. As sequências dos transcritos foram anotadas automaticamente utilizando a plataforma Egene (versão 2), onde 72% dos *contigs* foram anotados como proteínas hipotéticas, 20% como proteínas de função conhecida e 8% como proteínas hipotéticas conservadas. Dos *contigs* que apresentaram similaridade com proteínas de função conhecida ou proteínas hipotéticas conservadas, 74,2% eram similares a proteínas descritas para o Filo Apicomplexa. Após o mapeamento no genoma de *Eimeria maxima*, foi observado que cerca de 77% dos *contigs* ORESTES mapearam. Para avaliar o perfil de expressão gênica entre os diferentes estágios estudados, os transcritos obtidos foram submetidos à análise digital da expressão empregando o *Northern Digital*, sendo observado que 73,5% dos *contigs* estavam presentes em um único estágio, 20,5% em dois estágios e 6% em três estágios, indicando um perfil de expressão gênica diferencial entre as três fases da esporulação. Os perfis digitais obtidos foram submetidos a um agrupamento hierárquico juntamente com os transcritos de merozoítos de *E. maxima*, gerados pelo grupo. Foi observado que os oocistos agrupavam num clado, enquanto que merozoítos agruparam num clado externo. Estes resultados são compatíveis com a biologia do parasita. Para classificar os

contigs em diferencial ou não diferencialmente expressos foi realizada análise estatística empregando o programa Kemp. Cerca de 69% dos *contigs* foram classificados como não diferencialmente expressos e 31% como diferencialmente expressos, dos quais a maioria foi observada em somente um estágio da esporulação. Dos *contigs* diferencialmente expressos, 34% das proteínas de função conhecida estavam relacionadas ao metabolismo, 5% ao complexo apical e o restante a proteínas de diversas funções. Para validação dos dados obtidos por ORESTES, a transcrição de um pequeno conjunto de genes foi avaliada por RT-qPCR. Os perfis de expresssão obtidos pelas duas técnicas foram bastante condizentes entre si.

Palavras-chave: *Eimeria maxima*. Coccidiose. Oocistos. Esporulação. Expressão gênica. ORESTES.

ABSTRACT

Manha APS. Differential expression analysis of *Eimeria maxima* oocysts in three phases of sporulation. [Ph. D. thesis (Biology of Host Pathogen Relationship)]. São Paulo: Instituto de Ciências Biomédicas da Universidade de São Paulo; 2011.

Coccidiosis of the domestic fowl is caused by protozoan parasites of the genus *Eimeria*. *E. maxima* is one of the most pathogenic species and widely distributed worldwide. In order to obtain a better understanding of the molecular mechanisms involved in oocyst sporulation of *E. maxima*, the gene expression profiles of unsporulated, sporulating (sporoblast phase) and sporulated oocyst were studied using the ORESTES methodology. This method allows obtaining partial transcript sequences derived mainly from the central portion of the mRNA, facilitating the discovery of new genes. We generated 20,149 total readings and after pre-processing, 75% of the sequences were accepted. Subsequent clustering of sequences, 1207 *contigs* and 2,091 singlets were obtained. The sequences of the transcripts were annotated automatically using the platform Egene (version 2) where 72% of the *contigs* were annotated as hypothetical proteins, 20% as proteins of known function and 8% as conserved hypothetical proteins. Amongst *contigs* showing similarity to proteins of known function or conserved hypothetical proteins, 74.2% protein were similar to those described for the phylum Apicomplexa. After mapping the genome of *Eimeria maxima*, it was observed that about 77% of the *contigs* ORESTES were mapped. For the evaluation of the profile of gene expression among different stages, the transcripts were submitted to digital analysis of expression using Northern Digital, and we observed that 73.5% of the *contigs* were present in a single stage, 20.5% in two internships and 6% in three stages, indicating a differential gene expression profiles among the three stages of sporulation. The digital profiles were submitted to a hierarchical clustering together with the transcripts of merozoites from *E. maxima*. It was observed that all oocysts phases clustered in a clade, whereas merozoites clustered in a separated clade. These results are consistent with the biology of the parasite. For sorting the *contigs* into differential or non-differentially expressed, a statistical analysis was performed using the program Kemp. About 69% of the *contigs* were classified as not differentially expressed and 31% as differentially expressed, most of which was observed in only one stage of

sporulation. Of the differentially expressed *contigs*, 34% of proteins of known function were related to metabolism, 5% apical complex proteins and the remaining proteins have various functions. To validate the data obtained by ORESTES, transcription of a small set of genes were assessed by RT-qPCR. The profiles obtained by two expression technique were quite consistent among themselves.

Keywords: *Eimeria maxima*. Coccidiosis. Oocysts. Sporulation. Gene Expression. ORESTES.

LISTA DE FIGURAS

Figura 1- Esquema de forma invasiva de um organismo do Filo Apicomplexa (<i>Toxoplasma gondii</i>).....	24
Figura 2- Produção e exportação mundial de carne de frango. Número (em milhões de toneladas) referente ao ano de 2009.	31
Figura 3- Ciclo de vida de <i>Eimeria</i>	33
Figura 4- Estrutura típica de oocistos esporulados de <i>Eimeria</i> spp.....	34
Figura 5- Fotomicrografias de oocistos esporulados das sete espécies de <i>Eimeria</i> de galinha doméstica	35
Figura 6- Eventos presentes na esporulação de oocistos de <i>E. tenella</i>	37
Figura 7- <i>Eimeria maxima</i> : sítio da lesão intestinal, lesão macroscópica e oocisto .	40
Figura 8- Saída em modo texto gerado pelo programa Kemp.	78
Figura 9- Fotomicrografias de oocistos de <i>E. maxima</i> em diferentes fases da esporulação.	88
Figura 10- Resultado do ensaio de PCR <i>multiplex</i> para detecção das sete espécies de <i>Eimeria</i> de galinha doméstica.	89
Figura 11- Eletroforese em gel de agarose do RNA mensageiro purificado e tratado com DNase RQ1.....	90
Figura 12- Resultado do ensaio de PCR de controle de integridade do RNA mensageiro de <i>E. maxima</i>	91
Figura 13- Resultado do ensaio de PCR para a verificação de contaminação de RNA mensageiro com DNA genômico de <i>E. maxima</i>	92
Figura 14- Eletroforese em gel de agarose 1,0% mostrando os perfis de cDNA das bibliotecas Mx093 (1) e Mx094 (2) provenientes de oocistos esporulados de <i>Eimeria maxima</i>	93
Figura 15- Produto da reação de PCR de colônia a partir de sequências clonadas da biblioteca de oocistos não esporulados, MxOu02601.	93
Figura 16- Histograma do número de leituras em cada <i>contig</i> em relação à sua frequência nas bibliotecas de <i>Eimeria maxima</i>	96
Figura 17- Histograma de distribuição do tamanho dos <i>contigs</i> (em nucleotídeos) em relação à sua frequência nas bibliotecas de <i>Eimeria maxima</i>	97
Figura 18- Alinhamento das sequências nucleotídicas dos genes completos depositados no Genbank (barra de cor preta, mais larga) com as sequências ORESTES similares (barra de cor preta, mais fina) e ORFs preditas (barra de cor cinza).....	99
Figura 19- Distribuição dos <i>contigs</i> totais de <i>E maxima</i> quanto à anotação automática.....	102
Figura 20- - Diagrama de Venn mostrando a distribuição de produtos proteicos provenientes de ORESTES em relação ao Filo Apicomplexa, à classe	

	Coccidia, ao Gênero <i>Eimeria</i> e à espécie <i>E.maxima</i> , após buscas de similaridade por BLAST em banco de dados não redundante	103
Figura 21-	Distribuição dos produtos proteicos anotados automaticamente similares a proteínas do gênero <i>Eimeria</i>	104
Figura 22-	Gráfico acíclico de termos GO (ontologia gênica) das sequências proteicas codificadas pelos cDNAs reconstruídos a partir dos <i>contigs</i> de ORESTES de <i>E. maxima</i> para a categoria componentes celulares.....	105
Figura 23-	Gráfico acíclico de termos GO (ontologia gênica) das sequências proteicas codificadas pelos cDNAs reconstruídos a partir dos <i>contigs</i> obtidos de leituras de ORESTES de <i>E. maxima</i> para a categoria processos biológicos	106
Figura 24-	Gráfico acíclico de termos GO (ontologia gênica) das sequências proteicas codificadas pelos cDNAs reconstruídos a partir dos <i>contigs</i> obtidos de leituras de ORESTES de <i>E. maxima</i> para a categoria funções moleculares	107
Figura 25-	Diagrama de Venn mostrando a distribuição das sequências obtidas quanto à fase da esporulação	108
Figura 26-	Dendrograma obtido pelo método de distancia de Aitchison a partir dos perfis de expressão gênica dos transcritos ORESTES de oocistos não esporulado (Ou) oocistos parcialmente esporulados (Op), oocistos esporulados (Os) e merozoítos (Mz) de <i>E. maxima</i>	109
Figura 27-	Distribuição dos <i>contigs</i> anotados automaticamente	111
Figura 28-	Comparação dos perfis de expressão gênica de três fases da esporulação de oocistos de <i>Eimeria maxima</i>	135
Figura 29-	Valores de Ct (PCR <i>Cycle threshold</i>) para os transcritos candidatos a genes controle.....	154
Figura 30-	Expressão relativa (\log_2) obtida pela técnica de Real-time PCR (qRT-PCR) para os genes codificantes de Dinamina; Aminopeptidase; RON (proteína rombóide); Proteína da família BT1; PUP (proteína putativa não caracterizada); Transhidrogenase; Eimepsina; Proteína de desenvolvimento de esporozoíto (PDE); HSP 90 (proteína de choque termico); Mic2 (proteína de micronema 2); ANT (antígeno não transmembranar); Serpina e HP (proteína hipotética) em oocistos parcialmente esporulados e esporulados em relação aos oocistos não esporulados (nível basal, eixo x).	160

LISTA DE TABELAS

- Tabela 1-** Registro dos principais marcos para o estudo do transcriptoma44
- Tabela 2-** *Primers* testados em experimentos de PCR em tempo real para a escolha do gene controle (*housekeeping gene* constitutivamente expresso)81
- Tabela 3-** *Primers* utilizados para a validação da expressão gênica diferencial em experimentos de PCR em tempo real..... 82-83
- Tabela 4-** Pré-processamento das leituras ORESTES de *Eimeria maxima*.....94
- Tabela 5-** Número de eventos e redundância (em porcentagem) observados após a montagem das sequências ORESTES de oocistos não esporulados (Ou), oocistos parcialmente esporulados (Op) e oocistos esporulados (Os) de *Eimeria maxima*.....95
- Tabela 6-** Lista dos cDNAs completos depositados no Genbank que eram similares (*e value* 0,0) aos transcritos ORESTES de *Eimeria maxima*.....98
- Tabela 7-** Mapeamento dos *contigs* ORESTES no genoma de *Eimeria maxima* empregando diferentes parâmetros de alinhamento e de similaridade
.....100
- Tabela 8-** Distribuição dos transcritos ORESTES diferencialmente expressos em função dos perfis de expressão encontrados nas diferentes fases da esporulação: (Ou) oocistos não esporulados, (Op) oocistos parcialmente esporulados e (Os) oocistos esporulados110
- Tabela 9-** *Contigs* diferencialmente expressos entre os estágios da esporulação de *E.maxima* que apresentaram resultado positivo de BLAST113-118
- Tabela 10-** *Contigs* não diferencialmente expressos entre os estágios estudados que apresentaram resultado positivo de BLAST119-130
- Tabela 11-** Motivos proteicos identificados pelo Interpro para *contigs* diferencialmente expressos entre as diferentes fases da esporulação em *E.maxima* cujo produto foi classificado como proteína hipotética. Estão apresentados na tabela a identificação dos transcritos e os motivos proteicos..... 131
- Tabela 12-** Resultados positivos do programa RPS-BLAST para *contigs* diferencialmente expressos cujo produto foi classificado como proteína hipotética. Estão apresentados na tabela a identificação dos transcritos e os resultados da busca pelo programa RPS-BLAST..... 132-133
- Tabela 13-** Relação dos 100 *contigs* diferencialmente expressos entre as três fases da esporulação de *Eimeria maxima*, que eram mais abundantes .137-141
- Tabela 14-** *Contigs* diferencialmente expressos entre os 100 transcritos mais abundantes de *E. maxima* cujos produtos foram similares aos transcritos de *E. tenella* e *E. acervulina*..... 142-146
- Tabela 15-** Relação dos *contigs* diferencialmente expressos entre os 100 mais abundantes de *E. maxima* cujos perfis de expressão gênica foram

similares para pelo menos uma espécie, *E. tenella* ou *E. acervulina*147

- Tabela 16-** Proteínas hipotéticas presentes nos 100 *contigs* diferencialmente expressos mais abundantes em *E. maxima* submetidas a análise no programa Threader 149-151
- Tabela 17-** Relação dos *contigs* diferencialmente expressos entre as três fases da esporulação de *Eimeria maxima*, que foram avaliados por RT-qPCR ..154
- Tabela 18-** Comparação dos perfis de expressão em três fases da esporulação de oocistos dos genes selecionados para estudos de RT-qPCR de *E. maxima* com os transcritos similares de *E. acervulina* e *E. tenella*155
- Tabela 19-** Concentração de *primers* utilizada nos experimentos de PCR em tempo real. Estão apresentadas na tabela a identificação do transcrito de *E.maxima*, a descrição do produto e a concentração dos *primers*. Eficiência da reação obtida a partir das reações realizadas.....155
- Tabela 20-** Eficiência da reação obtida a partir das reações realizadas.....157
- Tabela 21-** Comparação da expressão gênica obtida pelas técnicas de qRT-PCR (RT- Real time PCR) e ORESTES para as fases da esporulação de *E.maxima* de maior e menor transcrição de 13 genes estudados.....160

LISTA DE ABREVIATURAS

- Água Milli-Q** - água purificada por osmose reversa, deionizada e tratada com luz UV (Sistema RiOs e Milli-Q Synthesis da Millipore)
- BLAST** - *Basic Local Alignment Search Tool* (Ferramenta de busca de alinhamento local básica)
- CAP3** - *Contig Assembly Program* (Programa de Montagem de *Contigs*)
- cDNA** - *complementary DNA* (DNA complementar)
- C-terminal** - carboxi-terminal
- Ct** - *Cycle Threshold* (o ponto *Ct* indica o momento a partir do qual o produto amplificado é quantificado).
- DEPC** - dietilpirocarbonato de sódio
- DNA** - ácido desoxirribonucleico
- DNase** - desoxirribonuclease
- dNTP** - abreviação coletiva para desoxinucleosídeos trifosfato (dATP, dCTP, dGTP e dTTP)
- DTT** - ditioneitol
- EDTA** - ácido etilenodiaminotetracético
- EST** - *Expressed Sequence Tags* (etiquetas de sequências expressas)
- GAM** - *gametocyte protein*
- G+C** - guanina + timina
- GO** - *Gene Ontology* (Ontologia Gênica)
- HTML** - *HyperText Markup Language* (Linguagem de Marcação de Hipertexto)
- HCl** - ácido clorídrico
- MOPS** - ácido 3-[N-morfilino] propanosulfônico
- NCBI** - *National Center for Biotechnology Information* (Centro Nacional de Informação em Biotecnologia - EUA)
- nr** - não redundante
- Op** - oocisto parcialmente esporulado
- ORESTES** - ORF ESTs (*Open Reading Frame Expressed Sequence Tags*)
- ORF** - *Open Reading Frame* (fase aberta de leitura)
- Os** - oocisto esporulado
- Ou** - oocisto não esporulado
- p/v** - peso por volume
- PCR** - *Polymerase Chain Reaction* (Reação em Cadeia da Polimerase)
- pH** - $-\log [H^+]$
- PBS** - Solução Salina Tamponada
- qPCR** - **PCR** - refere-se a tecnologia de PCR em tempo real (*Real Time PCR*)
- RAPD** - *Random Amplified Polymorphic DNA* (polimorfismo de DNA amplificado randomicamente)
- RNA** - ácido ribonucleico

RNA_m – RNA mensageiro

RNA_r – RNA Ribossômico

RNA_{se} – ribonuclease

RNA_t - RNA transportador

RT - *reverse transcription* - transcrição reversa

SCARs – *Sequence Characterized Amplified Regions* (Regiões amplificadas caracterizadas por Sequenciamento)

SDS – dodecil sulfato de sódio

SDS-PAGE - eletroforese em gel de poliacrilamida contendo SDS

spp. - espécie

TBE - solução tampão Tris-borato-EDTA

TE - solução tampão Tris-EDTA

TM - temperatura de fusão (de um *primer*)

TRF - *Tandem Repeats Finder* (Localizador de repetições seriadas)

Tris - Tris[hidroxmetil]aminometano

rRNA - RNA ribossômico

UV - ultravioleta

URL - *Uniform Resource Locator* (Localizador Uniforme de Recursos)

USDA - *United States Department of Agriculture* (Departamento de Agricultura dos Estados Unidos)

UTR - *Untranslated Region* (região não traduzida)

v/v - volume por volume

WFB - wall forming bodies

LISTA DE UNIDADES

°C - grau Celsius

cm - centímetro

g - aceleração da gravidade

g - grama

h - hora

kb - quilobase

kDa - quilodalton - 1.000 unidades de massa atômica (1 dalton = 1 u.m.a. = $1,66 \times 10^{-24}$ g)

L - litro

M - molar

mg - miligrama

mL - mililitro

mM - milimolar

N - normal

ng - nanogramas

nm - nanômetro

bp – (*base pair*) pares de bases

pmoles - picomoles

rpm - rotações por minuto

U - unidades

V - Volt

µg - micrograma

pg - picograma

fg - fentograma

µL - microlitro

µM -micromolar

SUMÁRIO

1 INTRODUÇÃO	23
1.1 O Filo Apicomplexa	24
1.2 O gênero <i>Eimeria</i>	25
1.3 Coccidiose aviária	25
1.3.1 Diagnóstico	26
1.3.2 Controle da doença	28
1.3.3 Importância econômica	30
1.3.4 Ciclo de vida	31
1.3.5 Oocistos e esporulação	33
1.4 <i>Eimeria maxima</i>	39
1.5 Caracterização do genoma de <i>Eimeria</i> spp	42
1.6 Transcriptoma	43
1.6.1 Metodologias utilizadas para a análise do transcriptoma	44
1.6.2 Transcriptoma de Apicomplexa	50
1.6.3 Transcriptoma de <i>Eimeria</i> spp	51
1.6.4 Transcriptoma de <i>Eimeria maxima</i>	54
2 OBJETIVOS	56
2.1 Objetivo geral	57
2.2 Objetivos específicos	57
3 MATERIAIS E MÉTODOS	58
3.1 Preparo e tratamento das soluções	59
3.1.1 Soluções	59
3.1.2 Meios de cultura	60
3.2 Condições de manutenção das aves	60
3.3 Propagação dos parasitas	61
3.3.1 Infecção das aves com oocistos de <i>Eimeria maxima</i>	61
3.3.2 Purificação dos oocistos	61
3.3.3 Esporulação de oocistos	62
3.4 Tratamento dos oocistos com hipoclorito de sódio	62

3.5 Verificação de contaminação interespecífica	63
3.6 Extração de RNA mensageiro de oocistos de <i>Eimeria maxima</i>	64
3.6.1 Verificação de contaminação do RNA mensageiro com DNA genômico	65
3.6.2 Verificação de integridade do RNA mensageiro	66
3.7 Construção de bibliotecas de cDNA do tipo ORESTES	67
3.7.1 Obtenção e manipulação dos perfis	67
3.7.2 Clonagem dos perfis.....	68
3.7.3 PCR de colônia	68
3.8 Reação de sequenciamento	69
3.9 Pré processamento das sequências	70
3.10 Agrupamento das sequências geradas por ORESTES	72
3.11 Verificação da posição relativa de transcritos ORESTES em cDNAs completos de <i>E. maxima</i>	73
3.12 Mapeamento dos <i>contigs</i> contra o genoma de <i>E. maxima</i>	73
3.13 Anotação automática das sequências geradas por ORESTES	73
3.14 <i>Northern</i> digital de sequências ORESTES de <i>E. maxima</i>	75
3.15 Agrupamento das sequências ORESTES utilizando o programa Simcluster	75
3.16 Análise estatística	76
3.17 Heat Map	78
3.18 Seleção e análise dos 100 reconstruídos diferencialmente expressos mais abundantes	79
3.18.1 Busca de similaridade de transcritos de <i>E. maxima</i> com sequências de outros parasitas do Filo Apicomplexa.....	79
3.18.2 Análise no programa Threader 3 (Protein Fold Recognition by Threading) 80	
3.19 PCR em tempo real	81
4 RESULTADOS	87
4.1 Propagação de parasitas	88
4.2 Verificação de contaminação interespecífica	88
4.3 Avaliação da qualidade das amostras de RNA mensageiro	89
4.4 Construção das bibliotecas ORESTES	92
4.4.1 Obtenção e manipulação dos perfis de cDNA	92
4.4.2 Clonagem, transformação bacteriana e PCR de colônia.....	93
4.5 Sequenciamento de ESTs do tipo ORESTES de <i>Eimeria maxima</i>	94
4.6 Pré-processamento das sequências ORESTES de <i>Eimeria maxima</i>	94

4.7 Agrupamento das sequências ORESTES de <i>E. maxima</i>	95
4.8 Análise da posição relativa dos transcritos ORESTES em cDNAs completos de <i>E. maxima</i>	97
4.9 Mapeamento dos <i>contigs</i> contra o genoma de <i>Eimeria maxima</i>	100
4.10 Anotação automática	101
4.11 Mapeamento dos termos de ontologia gênica (GO)	104
4.12 <i>Northern</i> Digital dos transcritos ORESTES de <i>E. maxima</i>	108
4.13 Análise dos dados no Programa Simcluster	108
4.14 Análise estatística	109
4.15 Análise comparativa da anotação automática de transcritos diferencialmente expressos com não diferencialmente expressos	111
4.16 <i>Heat Map</i>	134
4.17 Análise dos 100 transcritos diferencialmente expressos mais abundantes	136
4.17.1 Comparação dos produtos anotados de <i>E. maxima</i> com sequências de outros parasitas do Filo Apicomplexa.....	142
4.17.1.1 <i>Comparação dos produtos anotados de <i>E. maxima</i> com sequências ORESTES de <i>E. acervulina</i> e ORESTES/ESTs de <i>E. tenella</i></i>	142
4.17.2 Análise no programa Threader 3 (<i>Protein Fold Recognition by Threading</i>)....	148
4.18 PCR em tempo real	
4.18.1 Seleção de genes constitutivamente expressos (controle).....	152
4.18.2 Seleção de genes diferencialmente expressos	153
4.18.3 Expressão diferencial por RT-qPCR.....	156
5 DISCUSSÃO	161
5.1 Purificação dos Parasitas e extração de RNAm	163
5.2 Construção de bibliotecas de ORESTES	163
5.3 Anotação automática das sequências	167
5.4 Expressão gênica diferencial	169
6 CONCLUSÕES	186
REFERÊNCIAS	189

1 INTRODUÇÃO

1.1 O filo Apicomplexa

Os organismos do Filo Apicomplexa são protozoários que apresentam complexo apical na extremidade anterior das fases invasivas (Levine, 1973; Current et al., 1990; Tomley et al., 2001; Tomley e Soldati, 2001). As estruturas deste complexo estão associadas à motilidade, adesão e penetração dos parasitas nas células hospedeiras e formação do vacúolo parasitóforo (Current et al., 1990; Bumstead e Tomley, 2000; Morrissette e Sibley, 2002). Como se pode observar na Figura 1, o complexo apical é composto por um conóide, anéis polares, microtúbulos subpeliculares, roptrias, micronemas e grânulos densos (Dubremetz et al., 1998; Santos et al., 2009).

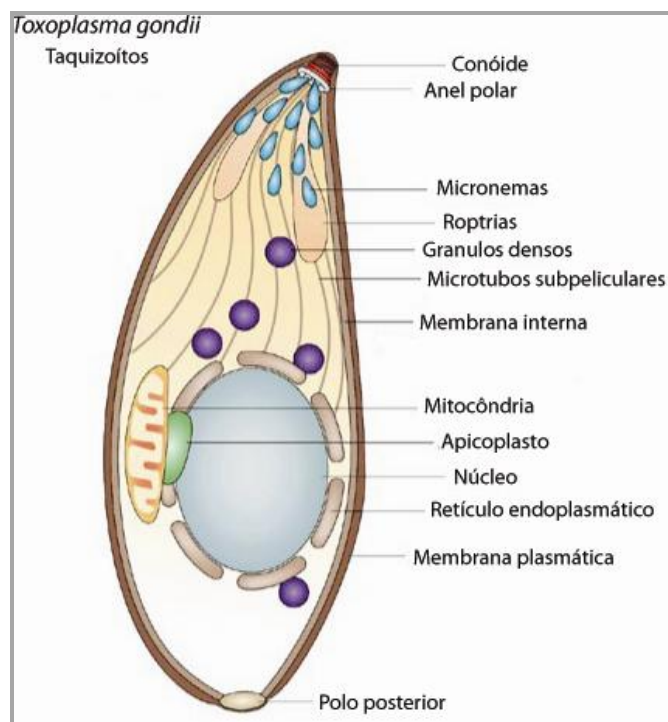


Figura 1- Esquema de forma invasiva de um organismo do Filo Apicomplexa (*Toxoplasma gondii*).

FONTE: Baum et al. (2006)

O Filo Apicomplexa é bastante diverso, sendo composto por mais de 5.000 espécies dentre os quais, patógenos de grande importância médica e veterinária,

tais como: *Toxoplasma gondii*, os gêneros *Plasmodium*, *Cryptosporidium*, *Cyclospora*, *Babesia*, *Neospora*, *Theileria* e *Eimeria* (Belli et al., 2005; Hu et al., 2006).

A maior parte das espécies de Coccídias conhecidas pertence à família Eimeriidae, compreendendo mais de 1700 espécies de protozoários (Levine, 1988) que podem infectar diversos hospedeiros vertebrados e invertebrados (Current et al., 1990).

1.2 O gênero *Eimeria*

O primeiro relato de protozoários do gênero *Eimeria* foi em 1674 por Antony van Leeuwenhoek após a visualização de oocistos de *Eimeria stiedai* na bile de um coelho. Entretanto, a primeira descrição destes parasitas só foi realizada cerca de 150 anos depois, quando Hake, em 1839, observou oocistos associados a um carcinoma no fígado de coelhos. Em 1875, Schineider classificou estes protozoários como pertencentes ao gênero *Eimeria* (Levine, 1973).

Parasitas do gênero *Eimeria* são os mais bem sucedidos da classe Coccidia, pois além de ter distribuição mundial e serem comumente observados (Levine, 1973), tem a capacidade de infectar grande variedade de hospedeiros vertebrados e invertebrados (Current et al., 1990). Estes protozoários possuem um ciclo de vida monoxênico com fases de reprodução assexuada e sexuada (Hammond e Long, 1973; Current et al., 1990). A merogonia e a gametogonia ocorrem dentro das células do hospedeiro e a esporogonia no meio ambiente (Levine, 1973).

1.3 Coccidiose aviária

Existem sete espécies de *Eimeria* que causam a coccidiose em galinha doméstica, são elas: *E. acervulina*, *E. brunetti*, *E. maxima*, *E. mitis*, *E. necatrix*, *E. praecox* e *E. tenella* (Fernando, 1990).

Cada espécie se multiplica em uma região específica do trato intestinal, ou seja, a infecção é sítio-específica. Por exemplo, *E. acervulina* (Tyzzer, 1929; Levine, 1961) e *E. praecox* infectam a porção superior do intestino delgado (Levine, 1961). *E. maxima* acomete a parte média do intestino, principalmente jejuno e íleo (Tyzzer, 1929; Levine, 1961). *E. necatrix* inicialmente infecta a porção média do intestino delgado durante as duas primeiras esquizogonias, já a terceira esquizogonia e a formação dos gametas ocorrem nos cecos (Levine, 1961). *E. tenella* infecta os cecos (Tyzzer, 1929). *E. mitis* acomete a porção inferior do intestino delgado (Tyzzer, 1929; Levine, 1961) e em *E. brunetti* a infecção inicial e a primeira esquizogonia ocorrem ao longo do intestino delgado, porém a segunda esquizogonia e formação gametas são observadas na porção posterior do intestino delgado e no intestino grosso até o reto (Levine, 1942; Levine, 1961).

Além de serem sítio-específicas, a patogenicidade deste protozoário também varia quanto à espécie, por exemplo, *E. tenella* e *E. necatrix* são as mais patogênicas, enquanto *E. mitis* provoca poucas lesões intestinais (Long e Joyner, 1984; Shirley, 1995). No campo, *E. tenella*, *E. maxima* e *E. acervulina* são as espécies mais relevantes e frequentes.

Estes parasitas acometem principalmente frangos de corte e matrizes reprodutoras (Williams, 1998), induzindo aumento da conversão alimentar, menor capacidade de absorção de nutrientes, menor ganho de peso e em alguns casos, aumento da mortalidade (Min et al., 2004).

As infecções comumente observadas no campo são mistas (Levine, 1961), entretanto, a resposta imune é espécie-específica, ou seja, uma ave imune a uma espécie de *Eimeria* é suscetível às demais (Dalloul e Lillehoj, 2005).

1.3.1 Diagnóstico

Classicamente, a identificação das espécies de *Eimeria* é realizada pela análise de diversos parâmetros, como: espécie do hospedeiro, sítios de infecção e desenvolvimento do parasita, aspecto das lesões macro e microscópicas, período de pré-patência, período mínimo de esporulação, morfologia do oocisto, entre outros (Long et al., 1976; Long e Joyner, 1984; Eckert et al., 1995). Entretanto, estas

características biológicas podem estar sobrepostas dificultando a diferenciação das espécies de *Eimeria* que acometem a galinha doméstica, além de necessitar de profissionais altamente capacitados.

Visando a discriminação destas espécies de *Eimeria*, vários ensaios moleculares foram desenvolvidos, dentre estes, os baseados na reação em cadeia da polimerase (PCR) que além de outras vantagens, podem ser realizados a partir de pequenas quantidades de DNA sendo mais sensíveis que as técnicas que até então eram empregadas como, análise eletroforética de isoenzimas (Shirley, 1975; Shirley, 1978; Shirley, 1997), separação eletroforética de proteína 2D (bidimensional) (Sutton et al., 1989) e diagnóstico baseado em hibridização com sondas de rRNA (Ellis e Bumstead, 1990).

A partir destes avanços, testes de alta sensibilidade e especificidade foram desenvolvidos, entretanto, nem todos os ensaios permitiam a detecção das sete espécies de *Eimeria*, como por exemplo, o direcionado para a região intergênica do gene da subunidade ribossômica 5S que era eficiente apenas para *E. tenella* (Stucki et al., 1993).

Outra técnica que foi empregada para o diagnóstico molecular de *Eimeria* de galinha doméstica é o RAPD (polimorfismo de DNA amplificado randomicamente). Este método é baseado em PCR e utiliza *primers* arbitrários que, em condições de baixa estringência, resultam em um perfil específico para cada espécie. Inicialmente, permitia somente o diagnóstico de seis espécies (*E. acervulina*, *E. brunetti*, *E. maxima*, *E. mitis*, *E. necatrix* e *E. tenella*) (Procunier et al., 1993). Anos mais tarde, o nosso grupo desenvolveu um teste de RAPD que é capaz de diferenciar as sete espécies de *Eimeria* de galinha doméstica (Fernandez et al., 2003a).

O primeiro ensaio molecular baseado em PCR que permitia a diferenciação das sete espécies de *Eimeria* foi direcionado para o ITS1 (espaçador interno transcrito 1) do *cistron* de RNA ribossômico (Schnitzler et al., 1998; Schnitzler et al., 1999), entretanto este teste não é capaz de identificar simultaneamente as diferentes espécies, sendo necessária a realização de sete reações.

Nosso grupo também desenvolveu um conjunto de marcadores SCAR (*Sequence Characterized Amplified Region*) específicos para as sete espécies de *Eimeria* (Fernandez et al., 2004). A partir destes marcadores, foi desenvolvido um ensaio de PCR multiplex que em uma única reação é capaz de detectar e identificar as sete espécies de *Eimeria* de galinha doméstica (Fernandez et al., 2003b).

Para a quantificação da carga parasitária foram desenvolvidos ensaios de PCR em tempo real, que além de detectar e discriminar as espécies, permitem a quantificação do nível de infecção. O primeiro teste, direcionado para o gene codificador da proteína de micronema 1 e sequências espécie-específicas de SCARs, era capaz de discriminar e quantificar somente quatro espécies de *Eimeria* (*E. acervulina*, *E. maxima*, *E. necatrix* e *E. tenella*) (Blake et al., 2008). Logo em seguida, foram desenvolvidos testes baseados em marcadores SCAR que eram capazes de identificar cinco espécies de *Eimeria* (*E. acervulina*, *E. brunetti*, *E. maxima*, *E. necatrix* e *E. tenella*) (Kawahara et al., 2008). Estudos mais recentes têm permitido a identificação, discriminação e quantificação das sete espécies de *Eimeria* que acometem a galinha doméstica (Vrba et al., 2010).

1.3.2 Controle da doença

Devido à alta produção de oocistos e a fácil disseminação destes parasitas entre as aves suscetíveis, o controle da coccidiose aviária é extremamente necessário, sendo realizado pelo uso de anticoccidianos e vacinas vivas.

Desde a década de 40, a quimioterapia profilática utilizando drogas anticoccidianas tem sido a principal forma de controle da coccidiose aviária (McDonald e Shirley, 2009). Devido a vários fatores, dentre os quais, resistência parasitária causada pelo uso indiscriminado dos quimioterápicos (Chapman, 2003) e preocupação do mercado consumidor com a possibilidade de contaminação da carne de frango com fármacos, o uso destas drogas tem decaído bastante nos últimos anos (Shirley et al., 2007).

Restrições legislativas à utilização destes fármacos têm crescido principalmente na Europa e Estados Unidos, inclusive, algumas drogas foram totalmente banidas. No Reino Unido, por exemplo, o número de anticoccidianos disponíveis comercialmente caiu de 17 em 2002, para apenas 10 em 2004 (Shirley et al., 2007).

Por outro lado, o emprego de vacinas vivas vem aumentando, em 2002 estavam disponíveis no mercado ou em fase de registro 13 vacinas de *Eimeria* de galinha (Chapman et al., 2002). Em áreas geográficas onde há resistência às

drogas, as vacinas vivas têm sido bastante utilizadas para o controle da coccidiose aviária (Williams, 1998; Allen e Fetterer, 2002).

A primeira vacina lançada no mercado foi a Coccivac® (Schering-Plough) em meados de 1950, sendo ainda empregada no campo (Chapman et al., 2002; McDonald e Shirley, 2009). A partir desta data, várias vacinas de *Eimeria* foram desenvolvidas, as quais podem ser classificadas em quatro tipos: vacinas vivas contendo cepas virulentas, vacinas vivas de cepas precoces, vacinas vivas de cepas atenuadas e as vacinas de DNA recombinante (Allen e Fetterer, 2002; McDonald e Shirley, 2009). Somente as vacinas vivas contendo cepas virulentas e precoces estão disponíveis para comercialização.

As vacinas compostas por cepas virulentas são formuladas com suspensões de oocistos viáveis oriundos de espécies patogênicas. É utilizada em baixas doses, o suficiente para gerar resposta imune sem causar lesões intestinais significativas (Shirley et al., 2005). Porém, erros na administração podem induzir a doença propriamente dita (Shirley et al., 2007).

Por apresentarem um menor número de ciclos de reprodução assexuada e conseqüentemente, redução do período de pré-patência, as vacinas de cepas precoces oferecem menos riscos de indução da doença (McDonald e Shirley, 2009). A desvantagem desta vacina está no seu custo de produção, já que o potencial reprodutivo destes parasitas é 90% menor que o das cepas originais.

Para obtenção de vacinas com cepas atenuadas, as cepas virulentas são submetidas a tratamento térmico ou irradiação (x ou γ) o que diminui a infectividade e a produção dos oocistos. Entretanto, estas cepas conferem somente resistência parcial às infecções (Allen e Fetterer, 2002; McDonald e Shirley, 2009).

Estudos para o desenvolvimento de vacinas recombinantes têm aumentado nos últimos 10 anos, entretanto, até o momento, nenhuma destas vacinas está disponível para comercialização. Um dos grandes obstáculos para utilização destas vacinas é a falta de proteção cruzada entre as espécies de *Eimeria* (Allen e Fetterer, 2002).

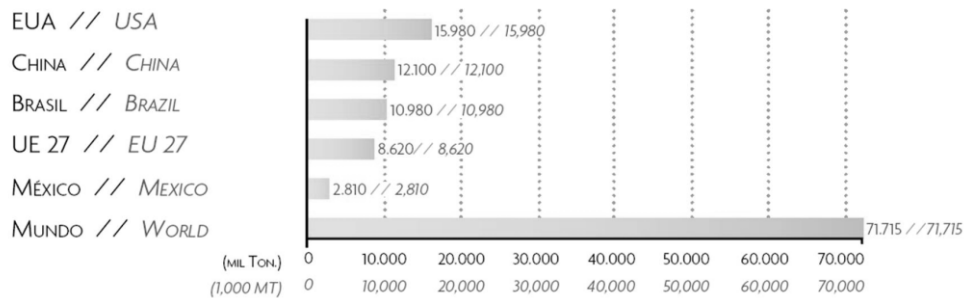
1.3.3 Importância econômica

Apesar de grandes avanços na pesquisa de quimioterápicos, vacinas e nutrição, a coccidiose aviária é umas das parasitoses mais importantes do ponto de vista econômico, causando grande prejuízo à indústria avícola (McDougald e Reid, 1995), seja, por custos diretos como menor ganho de peso, aumento de mortalidade, susceptibilidade a infecções secundárias e tratamento com quimioterápicos, ou por custos indiretos como uso de drogas e/ou vacinas para a prevenção da infecção (Shirley et al., 2005).

Calcula-se que os gastos mundiais relacionados ao controle desta doença variem de 800 milhões (Allen e Fetterer, 2002) a três bilhões de dólares por ano (Shirley et al., 2004b).

A indústria avícola tem um papel fundamental na economia brasileira. O Brasil é o primeiro maior exportador e o terceiro maior produtor de carne de frango do mundo (Figura 2). Em 2009 foram produzidas mais de 11 milhões de toneladas de carne de frango, das quais, 67% foram consumidas pelo mercado interno e a exportação deste produto resultou em uma receita cambial de quase 5,86 bilhões de dólares (Associação Brasileira dos Produtores e Exportadores de Frangos, Relatório Anual 2009-10 (http://www.abef.com.br/Relatorios_Anuais.php)).

PRODUÇÃO MUNDIAL DE CARNE DE FRANGO



EXPORTAÇÃO MUNDIAL DE CARNE DE FRANGO

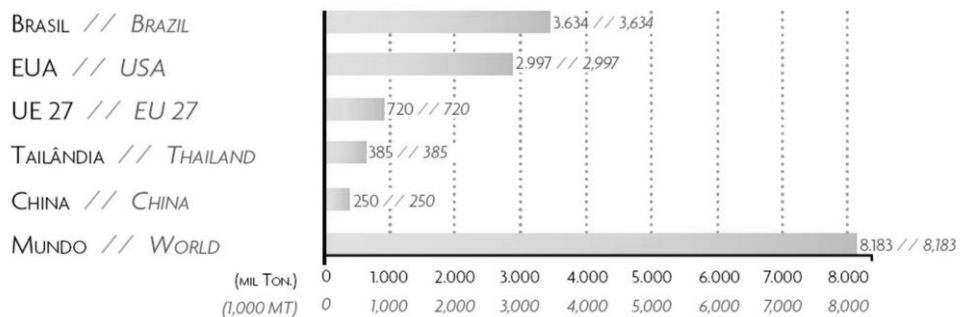


Figura 2- Produção e exportação mundial de carne de frango. Número (em milhões de toneladas) referente ao ano de 2009.
FONTE: ABEF (2009).

1.3.4 Ciclo de vida

O ciclo de vida de parasitas do gênero *Eimeria* está representado na Figura 3. A infecção das aves se inicia pela ingestão de um oocisto esporulado (Etapa 1). Em aves como a galinha doméstica (*Gallus gallus*), a parede do oocisto é rompida mecanicamente na moela por ação da musculatura e abrasão provocada por detritos ingeridos. Uma vez no intestino delgado, na presença de tripsina e sais biliares, ocorre a excitação, na qual os esporozoítos saem ativamente dos esporocistos (Etapa 2). Assim que liberados, os esporozoítos invadem as células epiteliais do intestino (Etapa 3), onde se multiplicam por fissão múltipla (reprodução assexuada) formando os esquizontes ou merontes (Etapa 4). Após o amadurecimento e ruptura dos esquizontes há liberação dos merozoítos (Etapa 5), os quais podem infectar

novas células intestinais por uma ou mais gerações, na dependência da espécie de *Eimeria* (Etapa 6). Após a esquizogonia, os parasitas se diferenciam em macrogametócito (Etapa 7) ou microgametócito (Etapa 8). Os microgametócitos se rompem e liberam os microgametas (Etapa 8), que por possuir dois flagelos são capazes de se locomover até os macrogametas (Etapa 9) fecundando-os e produzindo o zigoto (Etapa 10). O zigoto pode ser produto da fertilização de gametas da mesma cepa (autofertilização) ou de cepas diferentes (fertilização cruzada) (Shirley e Harvey, 1996). O zigoto maduro origina o oocisto (Etapa 10), único estágio diplóide, o qual é liberado no ambiente juntamente com as fezes sob a forma de oocisto não esporulado. Em condições adequadas de temperatura, umidade e oxigenação, o oocisto sofre uma meiose seguida de uma mitose formando os esporoblastos (Etapa 11). Cada esporoblasto sofre uma segunda mitose originando os esporocistos os quais contém dois esporozoítos (Etapa 1). O ciclo de vida é completado quando o oocisto esporulado é ingerido por um hospedeiro susceptível (Hammond e Long, 1973; McDougald e Reid, 1995; Min et al., 2004).

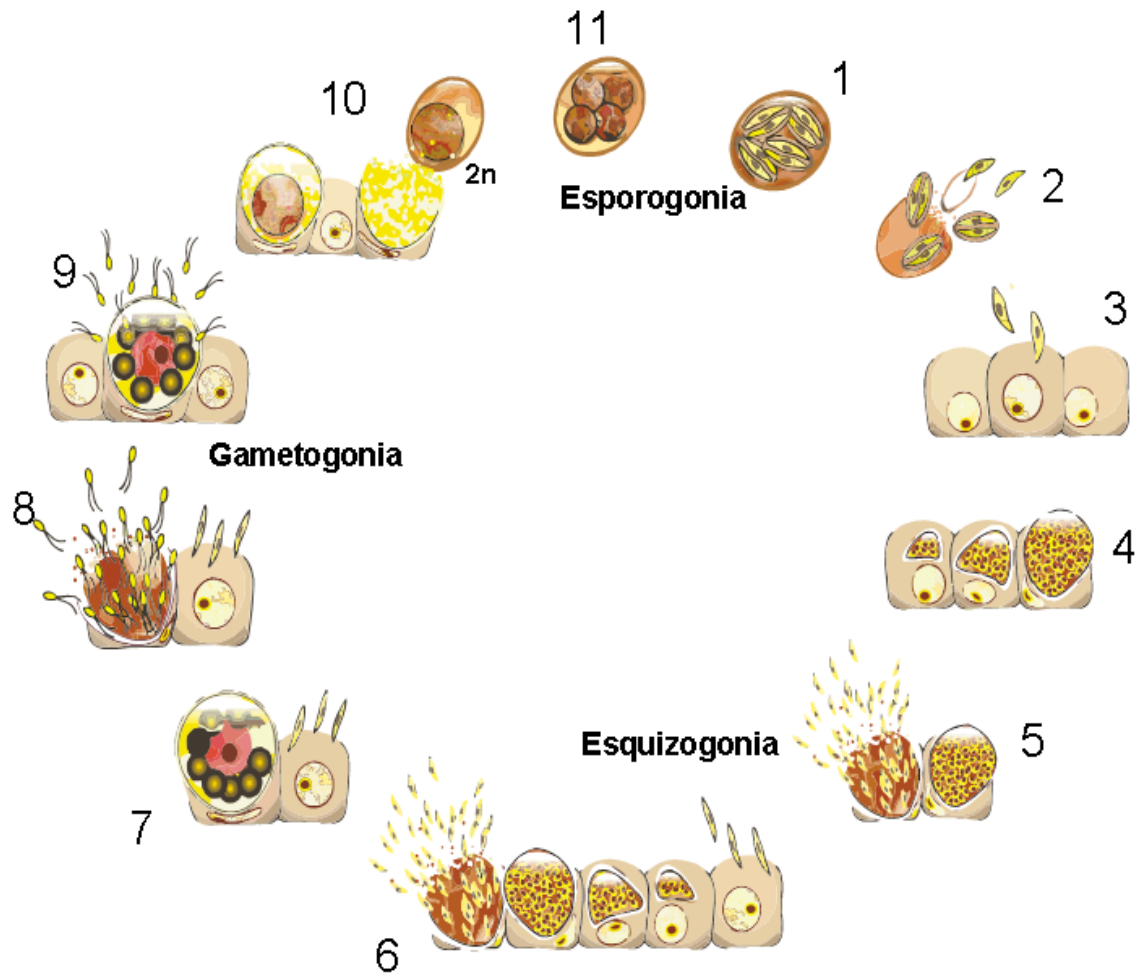


Figura 3- Ciclo de vida de *Eimeria*.

FONTE: Pôster de divulgação: "Entendendo e Controlando a Coccidiose aviária", Arte gráfica: Helton Barreiro, Patrocínio Biovet, Edição geral Oswaldo Gessulli Neto (Gessulli Agribusiness), com permissão.

1.3.5 Oocistos e esporulação

Por ser uma fase exógena do ciclo de vida dos Coccídios (Current et al., 1990), o oocisto é o estágio mais acessível para a realização de estudos biológicos. A morfologia típica de oocisto de *Eimeria* spp. é mostrada na Figura 4.

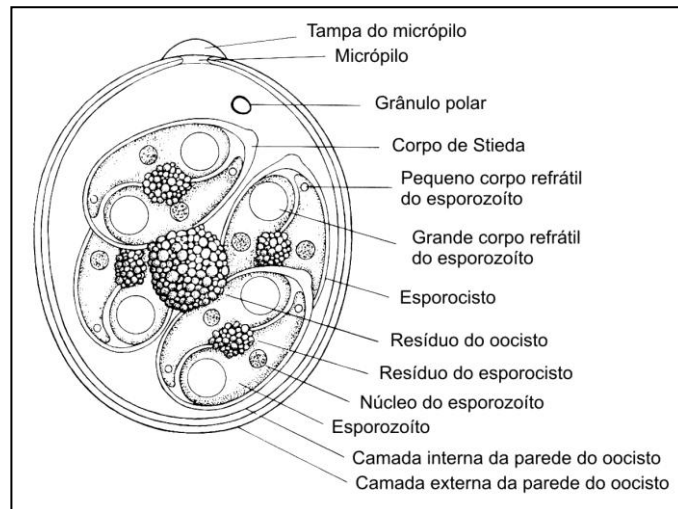


Figura 4- Estrutura típica de oocistos esporulados de *Eimeria* spp.
 FONTE: adaptado de Levine (1961).

A cor, o tamanho (área e diâmetro), o formato (elíptico, ovóide ou circular), a textura da superfície, a estrutura interna (presença ou ausência de grânulo polar, corpos residuais, corpúsculo de *Stieda* e micrópilo visível) e diversidade no formato dos esporocistos são características que podem variar de espécie para espécie (Joyner, 1982; Castañón et al., 2007). A Figura 5 apresenta fotomicrografias de oocistos das sete espécies de *Eimeria* que infectam a galinha doméstica. Como podemos observar os oocistos apresentam características morfológicas comuns e particulares de cada espécie. Por exemplo, os oocistos de *E. acervulina* têm tamanho similar ao de *E. mitis*, já oocistos de *E. maxima* são maiores que os demais. Em relação à forma, oocistos de *E. praecox* e *E. mitis* são circulares, os de *E. maxima* e *E. brunetti* tem forma ovóide e os de *E. acervulina* e *E. necatrix* são elípticos (Castañón, 2006).

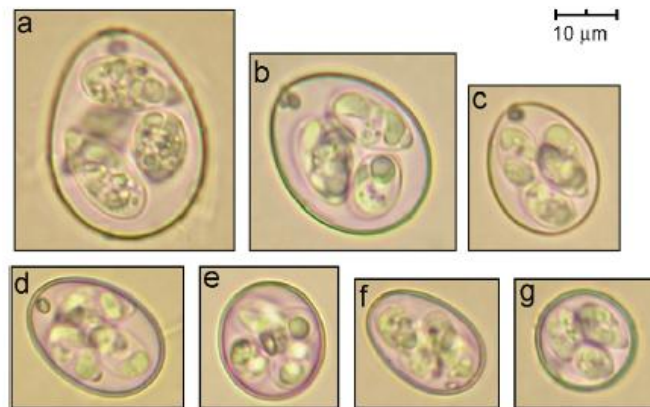


Figura 5- Fotomicrografias de oocistos esporulados das sete espécies de *Eimeria* de galinha doméstica. (a) *E. maxima*, (b) *E. brunetti*, (c) *E. tenella*, (d) *E. necatrix*, (e) *E. praecox*, (f) *E. acervulina* e (g) *E. mitis*.
 FONTE: Castañón (2006).

A parede do oocisto é extremamente robusta (Mai et al., 2009) sendo resistente a proteólise, além de ser impermeável a várias soluções como por exemplo, detergentes e desinfetantes. Entretanto, é permeável à amônia, ao bromometano, a determinadas moléculas e substâncias lipossolúveis (Monné e Hönig, 1954; Ryley, 1973; Mai et al., 2009).

A parede dos oocistos contém camadas multilamelares sendo composta por lipídeos, carboidratos e proteínas. Dentre os lipídeos, são descritos esteróides, ácidos graxos (docosanol, tetracosanol, hexacosanol, miristato, palmitato, estearato, oleato, linolato e fosfolipídios) e alcoóis graxos (Stotish et al., 1978). Em relação aos carboidratos, são observados manose, galactose, glicose e hexosamina (Stotish et al., 1978).

Há diversas proteínas envolvidas na formação da parede do oocisto, GAM56 e GAM82 são as mais estudadas e foram identificadas em oocistos de *E. acervulina*, *E. maxima* e *E. tenella*. Estas proteínas têm importante papel na maturação da parede do oocisto (Belli et al., 2009).

A parede dos oocistos é formada a partir do conteúdo de organelas encontradas exclusivamente em macrogametas, corpos específicos do tipo 1 (WFB1) e do tipo 2 (WFB2) (Pittilo e Ball, 1980), cujos tamanhos variam de acordo com a espécie de *Eimeria* (Mai et al., 2009).

Uma vez no ambiente, os oocistos passam por um processo denominado de esporulação ou esporogonia, tornando-se infectantes. Para tanto, são necessárias

condições ideais de temperatura, umidade e oxigenação (Fayer, 1980; Graat et al., 1994; Waldenstedt et al., 2001), sendo esta última, indispensável para o desencadeamento da esporulação (Wilson e Fairbairn, 1961).

O período de esporulação pode variar de espécie para espécie (Norton e Chard, 1983). Sob condições ideais (oxigenação, umidade relativa entre 40 e 80% e temperatura entre 28 °C a 30 °C) a esporulação é completada de um modo geral, entre 24 a 48 horas (Edgar, 1955; Fernando, 1990; Graat et al., 1994). Em *Eimeria maxima*, após 48h do início da esporulação já é possível observar oocistos totalmente esporulados (Edgar, 1955; Waldenstedt et al., 2001).

A esporulação é um processo aeróbico e o consumo de oxigênio varia ao longo da esporulação, principalmente durante a divisão nuclear (Wagenbach e Burns, 1969).

Uma grande quantidade de energia é consumida durante a esporulação e como o oocisto está em meio extracelular, esta é gerada à custa de carboidratos armazenados durante o ciclo intracelular (Edgar et al., 1944; Ryley, 1973; Wang et al., 1975). O acúmulo dos grânulos de polissacarídeos nos coccídeos ocorre durante a formação do zigoto e maturação do oocisto. Estima-se que em cada oocisto não esporulado são estocados aproximadamente 5×10^{-13} mol de resíduos de glicose (Wang, 1982). Durante a esporulação, inicialmente há utilização dos grânulos de polissacarídeos, seguida pelo metabolismo de lipídeos (Wilson e Fairbairn, 1961; Joyner, 1982). Em *E. acervulina*, estima-se que durante as primeiras 10 horas de esporulação o nível de polissacarídeo estocado decresce de 83 para 46 $\mu\text{g}/10^6$ oocistos (Wilson e Fairbairn, 1961).

O processo de esporulação inclui três divisões nucleares (uma meiose e duas mitoses) (Vetterling e Doran, 1969; Hammond, 1973; Ryley, 1973) e está exemplificado na Figura 6. O oocisto não esporulado apresenta um citoplasma que preenche totalmente o zigoto, o núcleo é arredondado e ocupa a região central do citoplasma (etapa 1). No primeiro dia da esporogonia o citoplasma dos oocistos condensa adquirindo o formato de uma massa esférica irregular, a qual é separada da parede do oocisto por um espaço preenchido por líquido (etapa 2). Logo em seguida, os grânulos citoplasmáticos agregam-se conferindo ao citoplasma uma aparência reticular (etapa 3). À medida que o citoplasma vai se contraíndo, o núcleo adquire um formato fusiforme e projeções nucleares podem ser visualizadas nas extremidades do eixo nuclear (etapa 4). O núcleo move-se para a região periférica

do oocisto (etapa 5) e ocorre a meiose (etapas 6 e 7). No início desta divisão os grânulos centriolares são visualizados no sentido oposto à membrana do núcleo, a membrana nuclear é rompida e o nucléolo desaparece (Canning e Anwar, 1968; Hammond, 1973). A seguir ocorre a primeira divisão mitótica (etapa 8), há citocinese formando quatro esporoblastos (etapas 9 a 10), geralmente esferoidais. Em seguida os esporoblastos se agrupam assumindo uma forma piramidal (etapa 11). Cada esporoblasto (etapa 12) irá sofrer uma segunda divisão mitótica, originando esporocisto com dois esporozoítos (etapa 13). Após esta etapa, o oocisto está totalmente esporulado.

A esporulação não é sincronizada (Vetterling e Doran, 1969; Hammond, 1973; Ryley, 1973) e nem todos os oocistos esporulam completamente. Em alguns oocistos não há a condensação do citoplasma (Levine, 1973), em outros, a esporulação cessa na fase de esporoblastos.

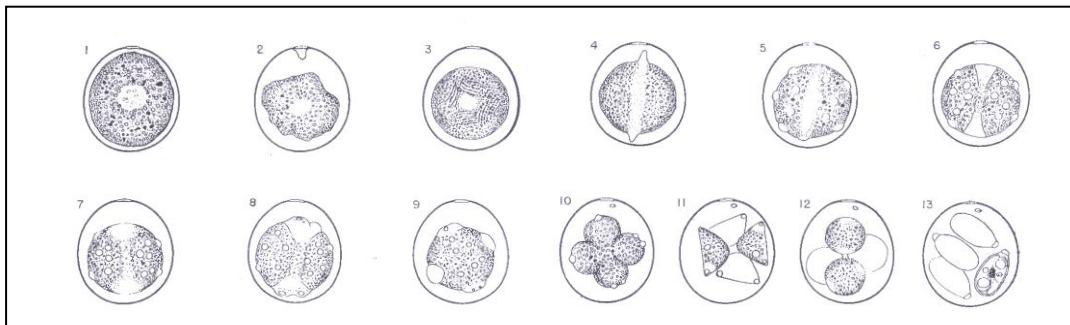


Figura 6- Eventos presentes na esporulação de oocistos de *E. tenella*. (1) oocisto não esporulado; (2) contração completa do citoplasma, forma irregular do citoplasma; (3) citoplasma reticulado; (4) projeção do eixo nuclear; (5) vesículas presentes na superfície citoplasmática e núcleo próximo à região periférica; (6) início da meiose; (7) divisão meiótica completada; (8) início da mitose; (9) início da citocinese; (10) e (11) citocinese completada; (11) aspecto piramidal dos esporoblastos; (12) esporoblastos maduros e (13) oocisto esporulado.

FONTE: adaptado de Wagenbach e Burns (1969).

Em *E. maxima*, após 3 horas do início da esporulação o citoplasma começa a contrair. Às 12 horas inicia-se a formação dos esporoblastos e entre 21 e 30 horas de esporulação há maturação do esporocisto e formação do esporozoíto (Pittilo e Ball, 1985).

As variações fenotípicas das amostras estão intimamente relacionadas à variabilidade genética decorrente da divisão meiótica que ocorre durante a esporulação. Esta divisão permite a recombinação dos cromossomos homólogos e a divisão reducional para formar os esporozoítos haplóides (Del Cacho et al., 2010).

Em *Eimeria* spp, os estudos relacionados a este processo são extremamente limitados, uma vez que a meiose ocorre exclusivamente nos oocistos (Canning e Anwar, 1968; Del Cacho et al., 2005). Inicialmente, acreditava-se que não ocorria a recombinação gênica, de forma que a meiose em *Eimeria* spp. poderia ser o resultado do pareamento dos cromossomos homólogos não duplicados e a variação individual era devido à combinação aleatória dos cromossomos (Canning e Anwar, 1968). Porém mais tarde, após o mapeamento de ligação gênica em *E. tenella* foi demonstrado que há recombinação durante a meiose (Shirley e Harvey, 2000; Shirley et al., 2004a).

Recentemente, descrições da dinâmica e comportamento dos cromossomos durante a prófase I da meiose de *E. tenella* têm sido realizadas, contribuindo para o entendimento da biologia do parasita (Del Cacho et al., 2008; Del Cacho et al., 2010).

Apesar dos oocistos serem fundamentais para a ocorrência da coccidiose aviária, os mecanismos moleculares envolvidos no processo de esporulação, bem como o perfil da expressão gênica durante este processo são pouco estudados e conhecidos. A esporogonia é uma interessante fase do ciclo de vida para o estudo da expressão gênica, pois além de ocorrer *in vitro*, sem a interferência do hospedeiro, pode ser realizada sob condições controladas, diferentes amostras podem ser retiradas e analisadas a qualquer momento da esporulação e além disso, RNAm e proteínas podem ser extraídos sem contaminação com debris ou material proveniente das células hospedeiras.

Estudos moleculares em *E. maxima* e *E. tenella* têm demonstrado variações na transcrição gênica ao longo da esporulação (Wang e Stotish, 1975; Ellis e Thurlby, 1991).

Até o momento, vários genes foram descritos em oocistos de *Eimeria* spp. nas diferentes fases da esporulação: MOP, proteína de oocisto (Fetterer e Barfield, 2003); eimepsina (Laurent et al., 1993; Jean et al., 2001a; Miska et al., 2004); profilina (Fetterer et al., 2004); proteína quinase dependente de ciclina de *E. tenella*, EtCRK2 (Kinnaird et al., 2004); aminopeptidase N (Miska et al., 2004; Fetterer et al.,

2005); serino proteases (Fetterer et al., 2007b); serpina (Miska et al., 2004; Fetterer et al., 2008); antígeno SO7 (Fetterer et al., 2007a); proteína associada ao corpo refrátil (Abrahamsen et al., 1994); proteína de choque térmico 70, HSP70 (Del Cacho et al., 2001); proteína de choque térmico 90, HSP90 (Clark et al., 1996; Miska et al., 2005; Péroval et al., 2006); proteínas de micronema (Ryan et al., 2000; Miska et al., 2004); proteínas rombóides; prolil endopeptidase; transhidrogenase; proteína quinase, (Miska et al., 2004); manitol-1-fosfato-desidrogenase (Schmatz, 1997; Allocco et al., 1999); manitol-fosfatase (Liberator et al., 1998), fator depolimerizante de actina; actina (Xu et al., 2008) e lactato desidrogenase (Schaap et al., 2004).

1.4 *Eimeria maxima*

Eimeria maxima é uma das espécies de *Eimeria* de galinha doméstica mais patogênicas que, além de ser comumente observada, apresenta ampla distribuição mundial (Schwarz et al., 2010). Foi descrita pela primeira vez por Tyzzer em 1929, quando estudava isolados do intestino delgado de galinhas domésticas com enterite.

Os parasitas da espécie *E. maxima* infectam a porção média do intestino delgado, principalmente jejuno e íleo. As infecções podem causar distensão das alças intestinais, má absorção de carotenóides e lesões do tipo petéquias na mucosa intestinal (Gregory, 1990; Shirley, 1995), Figura 7 A e B.

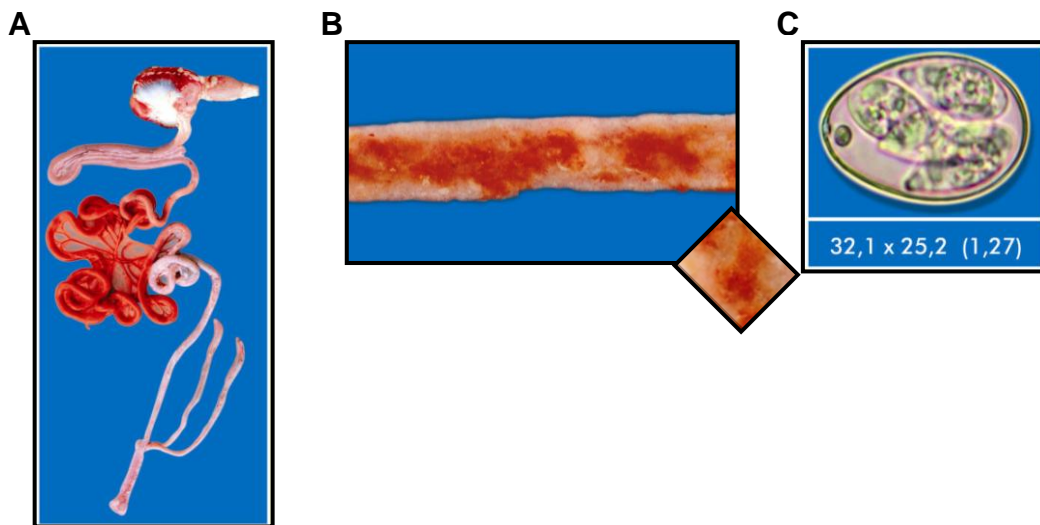


Figura 7- *Eimeria maxima*: sítio da lesão intestinal, lesão macroscópica e oocisto. (A) O intestino delgado, destacado em vermelho representa o sítio de colonização desta espécie, (B) lesões petequiais podem ser visualizadas na mucosa, (C) oocisto de *E. maxima*, com as medidas de tamanho referentes ao diâmetro maior e menor e a razão entre eles que está entre parênteses.
 FONTE: Pôster de divulgação: “Entendendo e Controlando a Coccidiose aviária”, Fotografia: Arthur Gruber. Patrocínio Biovet. Edição geral Oswaldo Gessulli Neto (Gessulli Agribusiness), com permissão.

A *E. maxima* apresenta grande variabilidade imunogênica entre cepas o que muitas vezes não confere proteção cruzada (Long e Millard, 1979; Fitz-Coy, 1992; Martin et al., 1997; Barta et al., 1998; Allen et al., 2005; Blake et al., 2005). Desta forma, para conferir uma boa proteção, algumas vacinas produzidas contra *Eimeria* contém mais de uma cepa de *E. maxima* (Long e Millard, 1979; Barta et al., 1998).

O ciclo de vida desta espécie foi primeiramente estudado por Tyzzer (1929). O período de pré-patência é de 5 a 6 dias. Após a infecção, os esporozoítos infectam as células epiteliais das vilosidades do duodeno e do íleo e após 4 a 5 dias, os esquizontes podem ser visualizados na mucosa intestinal. Os esquizontes localizam-se ao redor do núcleo da célula hospedeira, são pequenos (10 μm de comprimento por 8 μm de largura) e contém de 8 a 16 merozoítos. Após o rompimento dos esquizontes e lise celular há liberação dos merozoítos de primeira geração que infectam novas células intestinais. Em *E. maxima* há duas gerações de esquizontes. Os merozoítos de segunda geração invadem células intestinais da zona subepitelial e diferenciam-se em macro ou microgametócitos, os quais, são

encontrados abaixo do núcleo da célula hospedeira. Os microgametócitos medem entre 30 e 39 μm por 22 e 33 μm e contém um grande número de microgametas biflagelados. Os macrogametas são menores medindo em média 19 por 15 μm (Long, 1959). Após a fertilização dos macrogametas há a produção do zigoto e formação dos oocistos.

Muitos dos danos causados por *E. maxima* estão relacionados à uma localização mais profunda dos estágios sexuais (macrogametas e microgametas) na mucosa intestinal (Tyzzer, 1929; Brackett e Bliznick, 1950; Long, 1959; Levine, 1961).

Inicialmente, acreditava-se que o número máximo de oocistos produzidos para cada oocisto ingerido era de aproximadamente 12.000 (Brackett e Bliznick, 1950), porém mais tarde foi observado que a taxa de produção de oocistos varia em função da idade da ave, da quantidade e tempo de estocagem do inóculo (Long, 1959).

Os oocistos de *E. maxima* tem forma ovóide, parede lisa ou ligeiramente rugosa de coloração amarelada (Levine, 1961) e tamanho médio variando entre 30,86 a 32,1 μm de comprimento por 23,29 a 25,2 μm de largura (Tyzzer, 1929; Long, 1959; Castañón, 2006), Figura 7C. Os oocistos apresentam grânulos polares e os resíduos são ausentes. Os esporocistos são alongados, com tamanho médio entre 15 e 19 μm de comprimento por 8 a 9 μm de largura, apresentam corpúsculo de Stieda e não apresentam resíduos (Levine, 1961). Os esporozoítos têm tamanho médio de 19 μm de comprimento por 4 μm de largura (Long, 1959; Levine, 1961).

O processo de esporulação de oocistos ocorre em aproximadamente 72 horas quando incubados a 28 °C sob oxigenação constante. Interessante notar que alguns oocistos já são infectivos após 30 horas de incubação (Edgar, 1955; Long, 1959).

Os estudos relacionados à *Eimeria maxima* são importantes por quatro razões: ampla distribuição mundial, alta imunogenicidade (Rose e Long, 1962; Blake et al., 2005; Schwarz et al., 2010); altos níveis de variação imunogênica quando comparada a outras espécies de *Eimeria* spp. (Barta et al., 1998; Lew et al., 2003; Schwarz et al., 2009; Schwarz et al., 2010) e pouco se sabe sobre a expressão gênica em *E. maxima* (Smith et al., 1994; Witcombe et al., 2003; Schwarz et al., 2010).

1.5 Caracterização do genoma de *Eimeria* spp

O genoma nuclear de *Eimeria* spp. é haplóide durante a maior parte do ciclo de vida deste parasita, com exceção do oocisto não esporulado que é diplóide.

O cariótipo do genoma nuclear de *E. tenella* (Shirley e Harvey, 2000) e *E. maxima* (Blake et al., 2011), foram determinados por PFGE (*pulsed-field gel electrophoresis*, eletroforese de campo pulsado). O genoma de ambas as espécies apresenta uma complexidade entre 55 e 60 Mpb e está organizado em 14 cromossomos com tamanho variando entre 1 e 7 Mpb em *E. tenella* (Shirley e Harvey, 2000) e 1,9 Mpb a >6 Mpb em *E. maxima* (Blake et al., 2011). Em *E. tenella* o conteúdo GC é de 53% (Shirley, 2000).

O genoma da cepa H de *E. tenella* foi sequenciado empregando a abordagem *shotgun* com uma cobertura de 8,4 vezes (Ling et al., 2007). Os dados estão publicamente disponíveis no sítio do Instituto Sanger (<http://www.sanger.ac.uk/resources/downloads/protozoa/eimeria-tenella.html>). O genoma de *E. maxima* foi sequenciado com cobertura de 12 vezes e aproximadamente 75% do genoma está predito. Os dados estão disponíveis no sítio do Instituto GENOMalaysia (<http://www.genomemalaysia.gov.my/emaxdb/>)

O genoma de *E. tenella* contém repetições de trinucleotídeos GCA/TGC que podem estar presentes em sequências codificadoras e não codificadoras. A importância destes trinucleotídeos ainda não é conhecida (Shirley, 1994; Shirley, 2000; Shirley et al., 2004b). Outra repetição comum é o heptâmero TTTAGGG encontrado frequentemente em repetições seriadas de regiões não codificantes. Apesar de ser semelhante a uma sequência telomérica, não há evidências de que estas sequências estejam nos telômeros dos cromossomos de *Eimeria* spp. (Shirley, 2000; Shirley et al., 2004b).

Em paralelo, os cromossomos 1 e 2 de *E. tenella* foram totalmente sequenciados e apresentam uma complexidade de 1Mpb e 1,2Mpb respectivamente (Ling et al., 2007). O cromossomo 1 possui uma organização genômica incomum segmentada em regiões R (*feature-rich*), ricas em repetições, apresentando elementos similares a transposons e repetições teloméricas, confirmando estudos preliminares que apontavam para uma distribuição frequente de regiões repetitivas no genoma de *Eimeria tenella* (Shirley, 2000; Shirley et al., 2004a). Também

apresenta segmentos livre de repetições P (*feature-poor*) (Ling et al., 2007). Ambos os segmentos P e R podem ser identificados no genoma de *E. maxima*, embora sejam menos distintos do que os encontrados no cromossomo 1 de *E. tenella* (Blake et al., 2011).

A análise do mapa de ligação gênica em *E. maxima* mostrou que a distribuição das repetições dos trinucleotídeos CAG e seus isômeros associados é altamente relacionada aos segmentos R. Nestes segmentos o conteúdo AT é maior e há um maior nível de variação das repetições (Blake et al., 2011).

Além do genoma nuclear, estão presentes nestes parasitas dois genomas extracromossômicos: o mitocondrial e o do apicoplasto ou plastídeo.

O genoma mitocondrial tem alta composição de bases A/T, cerca de 65%, é constituído de concatâmeros lineares contendo unidades repetitivas de 6kb, sendo composto apenas de três genes estruturais: citocromo b, citocromo oxidase 1 e citocromo oxidase 3 e oito fragmentos de RNA ribossômico, os quais apresentam-se truncados e espalhados ao longo do genoma (Chapman e Shirley, 2003; Romano, 2004).

O apicoplasto, organela exclusiva dos organismos do Filo Apicomplexa, possui um genoma circular de aproximadamente 35 kb (Cai et al., 2003) e tem alta composição de bases A/T, cerca de 79,4% (Cai et al., 2003). O genoma contém entre 30 e 50 genes constitutivamente expressos que são geralmente relacionados à transcrição e tradução (Wilson et al., 1996; Fleige et al., 2010).

1.6 Transcriptoma

Há muitas décadas, o estudo da correlação entre a função gênica e seus padrões de expressão têm sido de grande interesse para diferentes ramos da biologia (Morozova et al., 2009).

Um dos pontos mais intrigantes da biologia molecular está relacionado ao fato de que células com mesmo mapa genético conseguem desempenhar funções distintas. Esta diversidade fenotípica tem sido associada à expressão de conjuntos de genes de diferentes funções (Morozova et al., 2009).

Os primeiros estudos de expressão gênica eram realizados em várias etapas, as quais compreendiam identificação das unidades transcricionais, mensuração dos níveis de expressão, estudo da diversidade transcricional (isoformas, transcrição alternativa, códon de iniciação e sítios de poliadenilação) e os fatores envolvidos na transcrição (Liu, 2005).

O advento de novas técnicas de estudo da expressão gênica iniciou uma nova era na biologia, pois ao invés de avaliar os genes individualmente, estas técnicas analisam o transcriptoma como um todo, o que tem permitido conhecer as alterações gênicas que ocorrem em diferentes condições biológicas ou mesmo patológicas, como os mecanismos moleculares envolvidos na relação entre o patógeno e o hospedeiro (Brady et al., 2006).

1.6.1 Metodologias utilizadas para a análise do transcriptoma

Os métodos para análise da expressão gênica evoluíram rapidamente em poucos anos (Morozova et al., 2009), Tabela 1.

Tabela 1- Registro dos principais marcos para o estudo do transcriptoma

Ano	Registro
1965	Primeiro sequenciamento da molécula de RNA
1977	Desenvolvimento da técnica de <i>Northern blot</i> e do método de sequenciamento de Sanger.
1989	Relatos da técnica RT-PCR para análise do transcriptoma
1991	Sequenciamento de EST em larga escala
1992	Introdução da técnica de <i>Differential Display</i> (DD) para a descoberta de genes diferencialmente expressos.
1995	Relatos das técnicas de <i>microarray</i> e de SAGE (<i>Serial Analysis of Gene Expression</i>)
2005	Primeira técnica de sequenciamento de nova geração (454/Roche) disponível no mercado
2006	Primeiro estudo de sequenciamento do transcriptoma utilizando sequenciadores de nova geração (454/Roche)

Fonte: Adaptada de Morozova et al. (2009).

O primeiro estudo do transcriptoma em larga escala foi publicado em 1991. Este trabalho avaliou um subconjunto de ESTs (*Expressed Sequence Tags*) provenientes de genes expressos de cérebro humano (Adams et al., 1991). Desde então, milhares de sequências de ESTs têm sido geradas para diferentes células, tecidos ou estágios de diversos organismos (Nagaraj et al., 2007). Estes estudos têm contribuído significativamente para a identificação de novos genes, determinação de regiões codificadoras de sequências genômicas, mapeamento físico, bem como análise do perfil de expressão gênica (Sterky e Lundeberg, 2000).

Para a construção das bibliotecas de ESTs, os cDNA são selecionados aleatoriamente e submetidos a um sequenciamento de uma única fita, gerando sequências curtas de aproximadamente 100-700 pb que podem ser obtidas a partir de um tecido ou estágio específico de desenvolvimento de um organismo. Para a construção de bibliotecas convencionais, as ESTs são obtidas a partir da região 5' ou 3' dos transcritos, utilizando-se adaptadores que permitem a clonagem unidirecional. Em bibliotecas direcionais 5', o sequenciamento apresenta maior cobertura de regiões codificadoras dos transcritos. Entretanto, as construções das bibliotecas direcionais 3' geralmente resultam na identificação das porções não traduzidas que apresentam maior variabilidade do que as regiões codificadoras, permitindo a identificação inequívoca destes transcritos, facilitando a obtenção de um perfil de transcrição quantitativo (Gruber, 2007).

Outra técnica também utilizada é a construção de bibliotecas de hibridização subtrativa. Esta metodologia permite reduzir a representatividade de transcritos já estudados, enriquecendo o conhecimento de padrões de expressão gênica de diferentes células, tecidos ou estágios (Gruber, 2007). Foi primeiramente utilizada na década de 80, para estudar a expressão gênica em gástrula de *Xenopus laevis* (Sargent e Dawid, 1983; Moody, 2001). Alguns anos mais tarde, esta técnica foi aprimorada sendo denominada de *Supression Subtractive Hybridization* (SSH). Além da subtração, esta técnica também permite a normalização da biblioteca (Diatchenko et al., 1996). Para a construção desta biblioteca os genes diferencialmente expressos são isolados após a hibridização de duas bibliotecas de cDNAs: uma amostra em excesso, ou seja amostra alvo da subtração (teste) e uma segunda amostra, que é a amostra controle. Os fragmentos semelhantes anelam-se, enquanto os diferencialmente expressos permanecem em fita única (*single-stranded*), podendo ser identificados (Gruber, 2007).

A geração de *fingerprinting* de RNA também foi uma das primeiras abordagens de estudo de genes diferencialmente expressos, incluindo o *Differential Display* (DD) (Liang e Pardee, 1992) e o RNA *fingerprinting by arbitrary primed* PCR (RAP_PCR) (Welsh et al., 1992). Ambos os métodos são baseadas na amplificação aleatória de subconjuntos de genes a partir de duas ou mais amostras de RNA (Moody, 2001).

O DD (*Differential Display*) é um método baseado na transcrição do RNAm utilizando oligo dT *primers*, *primers* ancorados ou *primers* arbitrários. Os cDNAs das diferentes amostras são separados em gel de poliacrilamida, clonados e sequenciados. A intensidade das bandas indica o nível de transcrição gênica. Além de identificar os genes diferencialmente expressos nos tecidos, células ou estágios, esta metodologia permite a comparação simultânea de várias amostras. Entretanto, este método não é quantitativo e falsos positivos podem ocorrer (Moody, 2001).

A avaliação da expressão gênica utilizando dados de *microarray* foi originalmente descrita da década de 90 em um estudo de transcriptoma de *Arabidopsis thaliana* (Schena et al., 1995). Um *microarray* é tipicamente definido como um arranjo pré-definido de moléculas ou fragmentos de DNA, cDNAs ou oligonucleotídeos quimicamente ligadas por ligações covalentes a uma superfície sólida ou à uma membrana (carregada positivamente). Este método baseia-se na hibridização entre as sondas que compõem estes microarranjos com o gene ou molécula de interesse. A ligação dos alvos marcados com fluoróforos permite a detecção e quantificação relativa das moléculas hibridizadas (Dufva, 2009). A expressão relativa de cada gene representado no *microarray* é avaliada pela comparação da intensidade de sinais de hibridização entre as amostras controle e experimental. Um dos aspectos mais interessantes é possibilidade de estudo de um grande conjunto de genes (3000 a 10000 genes) (Moody, 2001), no entanto, é necessário um conhecimento prévio das sequências a serem estudadas (Green et al., 2001). Por se tratar de um método de hibridização são necessárias replicatas experimentais, hibridizações cruzadas podem ocorrer e os dados de fluorescência somente detectam a presença ou ausência das moléculas hibridizadas (Dufva, 2009). Desta forma, a escolha correta das sondas é de fundamental importância para o sucesso do experimento, bem como, a realização de complexa análise estatística para a correta interpretação dos dados (Dufva, 2009). Os *microarrays* mais conhecidos são os de cDNA (Schena et al., 1995), de oligonucleotídeos (Pease

et al., 1994) e os *Affymetrix chips*, que são lâminas para hibridização comercialmente disponíveis (www.affymetrix.com/index.affx).

Outros métodos de análise de expressão gênica construídos a partir de pequenas sequências (*tags*) foram também desenvolvidos incluindo, SAGE (*Serial Analysis of Gene Expression*), MPSS (*Massive Parallel Signature Sequencing*), PETs (*Paired-end diTagging*), CAGE (*Cap Analysis Gene Expression*), RNA-seq e ORESTES (*Open Reading Frame Expressed Sequence Tags*).

SAGE (*Serial Analysis of Gene Expression*) e suas variações são considerados métodos de perfil de expressão digital, pois os dados de expressão gênica obtidos são analisados utilizando-se a contagem dos transcritos. A expressão dos transcritos é medida em termos absolutos, permitindo uma análise global da expressão gênica. Esta técnica foi criada na década de 90 (Velculescu et al., 1995) com a finalidade de obter perfis de expressão gênica e baseia-se no conceito de que pequenas *tags* são suficientes para identificar a transcrição de diferentes genes. Cada *tag* é gerada a partir da transcrição de RNAm, por uma digestão enzimática em um local pré-definido (Moody, 2001), logo após, as *tags* são concatenadas, sequenciadas e quantificadas (Tuteja e Tuteja, 2004). Cada *tag* representa uma sequência expressa e a contagem da mesma é diretamente proporcional à sua frequência em uma dada população de RNAm (Dinel et al., 2005). Uma das vantagens desta técnica é a possibilidade de gerar dados quantitativos (Moody, 2001), além de permitir a descoberta de novos transcritos (Rivals et al., 2007) e transcritos de baixa abundância (Chen et al., 2002; Kim et al., 2006). Entretanto, por gerar sequências curtas (entre 14 e 21pb), a identificação das *tags* e montagem dos transcritos pode ser prejudicada (Morozova et al., 2009). Além disso, devido à utilização de enzimas específicas para gerar as bibliotecas, alguns transcritos podem não estar representados (Moody, 2001).

Similar à técnica de SAGE, a metodologia MPSS (*Massive Parallel Signature Sequencing*) (Brenner et al., 2000) é baseada na produção de sequências curtas (16-20 bases) adjacentes ao sítio de restrição da enzima *DpnII* mais próximo à extremidade 3' do RNAm (Liu, 2005). Este método combina clonagem e amplificação de cDNA na superfície de milhares *microbeads* à uma alta capacidade de sequenciamento de DNA não baseado em gel. Cada corrida pode gerar cerca de 250.000 – 400.000 sequências. Como esta tecnologia é baseada em instrumentos sofisticados e protegida por direitos autorais, seu custo é elevado, sendo disponível

apenas comercialmente, inicialmente pela *Lynxgengen Therapeutica, Inc.* e, posteriormente pela *Solexa Inc.* que encerrou esta prestação de serviço em 2006.

Outra abordagem similar as descritas anteriormente é o CAGE (*Cap Analysis Gene Expression*), sendo descrita pela primeira vez em 2003 (Shiraki et al., 2003). Esta técnica é capaz de mensurar o nível de expressão gênica em larga escala pelo isolamento e sequenciamento de *tags* geradas a partir da extremidade 5' dos transcritos de RNAm (Hoon e Hayashizaki, 2008). Além disso, este método também possibilita a identificação dos promotores responsáveis pela transcrição do RNAm (Shiraki et al., 2003).

Na metodologia PET (*Paired-End diTagging*), *tags* pareadas são extraídas das extremidades dos fragmentos do cDNA, clonadas e sequenciadas em larga escala (Fullwood et al., 2009). A avaliação da expressão gênica é feita de forma digital, onde cada *tag* representa uma sequência expressa. Como as *tags* são pareadas, os transcritos obtidos podem ser mais facilmente mapeados (Ng et al., 2007).

A técnica de RNAseq é baseada no sequenciamento massivo de bibliotecas de pequenos ESTs, sendo possível gerar milhões de sequências de 30 a 400 nucleotídeos, dependendo da tecnologia de sequenciamento utilizada (Wang et al., 2009). Para tanto, uma população de RNA (total ou poliA+) é convertida numa biblioteca de fragmentos de cDNA empregando adaptadores que se ligam à uma ou ambas as extremidades. Cada molécula com ou sem etapa de amplificação é sequenciada em sequenciadores de última geração de forma a se obter pequenas sequências de uma extremidade (*single-end sequencing*) ou das duas extremidades (*pair-end sequencing*). As leituras são mapeadas contra o genoma ou transcriptoma e o nível de transcrição pode ser deduzido a partir do número de leituras encontradas para cada gene (Wang et al., 2009). Porém, nesta metodologia, as informações sobre o direcionamento das fitas podem se perdidas prejudicando o mapeamento das sequências geradas, principalmente quando os transcritos *sense* e *antisense* estão sobrepostos. Além disso, para o correto mapeamento das leituras é necessário um conhecimento prévio do genoma ou do transcriptoma (Cloonan e Grimmond, 2008).

Outra metodologia, também, aplicável para obtenção de um transcriptoma é o ORESTES (*Open Reading Frame ESTs*) a qual foi utilizada neste trabalho.

A técnica de ORESTES (*Open Reading Frame ESTs*) baseia-se na construção de mini bibliotecas utilizando *primers* arbitrários e PCR de baixa estringência para amplificação dos cDNA, gerando perfis de expressão gênica (Dias Neto et al., 2000). Como cada *primer* atua tanto como *primer forward* e *reverse*, há uma maior cobertura da região central da sequência dos RNAs mensageiros, a qual contém a maior parte da porção codificadora dos genes. Como estas bibliotecas são normalizadas, aumenta-se a probabilidade de detecção de transcritos raros (Gruber, 2007).

O primeiro trabalho utilizando esta metodologia foi realizado em 2000. Foram geradas 10.000 sequências do tipo ORESTES provenientes de tumor de mama humano. Além de contribuir com a descoberta de novas sequências, os resultados mostraram que a maior parte (71%) dos transcritos era de baixa abundância sendo encontrados preferencialmente na região codificante do gene (Dias Neto et al., 2000).

Após este estudo, diversos trabalhos visando a detecção de novos genes ou o estudo da expressão gênica diferencial em diversos tecidos (normais ou tumorais) ou em células de humanos foram realizados. Dentre estes podemos citar a identificação de genes presentes no cromossomo 22 de células provenientes de vários tumores humanos (de Souza et al., 2000); estudos comparativos da expressão gênica em células normais e tumorais da próstata (Sarkis, 2000; Fonseca et al., 2006), de mama (Leerkes et al., 2002; Nagai et al., 2004; Fonseca et al., 2006), da tireóide (Arnaldi et al., 2005) e de diferentes tecidos (Camargo et al., 2001; Brentani et al., 2003; Mello et al., 2009); identificação de genes relacionados à síndrome tricorinofalangiana do tipo I em pacientes com câncer de mama (Radvanyi et al., 2005); estudos comparativos da expressão gênica obtida por ESTs e ORESTES em células humanas (Sakabe et al., 2003), bem como identificação em larga escala de genes presentes em queratinócitos granulares humanos (Toulza et al., 2007).

Estudos do transcriptoma utilizando a metodologia ORESTES foram também realizados para outros organismos, como *Bos indicus* e *Bos taurus* (da Mota et al., 2004); *Apis mellifera* (Nunes et al., 2004); *Drosophila melanogaster* (Maia et al., 2007), em ratos como modelo de estudo da epilepsia (Avedissian et al., 2007) e em *Biomphalaria glabrata* (Lockyer et al., 2007; Hanelt et al., 2008; Lockyer et al., 2008).

A expressão gênica empregando ORESTES também foi estudada em larga escala para parasitas como *Paracoccidioides brasiliensis* (Garcia et al., 2010); *Trypanosoma rangeli* (Snoeiijer et al., 2004; Grisard et al., 2010); *Taenia solium* (Almeida et al., 2009) e *Schistosoma mansoni* (Fietto et al., 2002; Verjovski-Almeida et al., 2003).

1.6.2 Transcriptoma de Apicomplexa

Há uma grande quantidade de estudos do transcriptoma de diferentes estágios evolutivos de parasitas do filo Apicomplexa (Wastling et al., 2009).

Os bancos de dados dbEST (Boguski et al., 1993), EuPathDB (Aurrecoechea et al., 2007), NCBI dbEST, ApiEST-DB (Li et al., 2004b) e Full-Parasires (<http://fullmal.hgc.jp/>) (Tuda et al., 2010) tem uma grande quantidade de sequências ESTs relacionadas à expressão gênica de parasitas do Filo Apicomplexa, dentre estes, *Babesia bovis*, *Cryptosporidium parvum*, *Cryptosporidium muris*, *Neospora caninum*, *Plasmodium berguei*, *Plasmodium falciparum*, *Plasmodium yoelli*, *Plasmodium vivax*, *Sarcocystis neurona*, *Toxoplasma gondii*, *Eimeria acervulina*, *Eimeria brunetti*, *Eimeria maxima* e *Eimeria tenella*.

A análise do transcriptoma de parasitas do Filo Apicomplexa tem sido realizada empregando várias metodologias, tais como SAGE (*Serial Analysis of Gene Expression*) para *P. falciparum* (Munasinghe et al., 2001), *T. gondii* (Radke et al., 2005) e *Eimeria tenella* (Novaes, 2009); *microarray* para *P. falciparum*, *P. berguei* e *T. gondii* (EuPathDB - <http://eupathdb.org>) e MPSS (*Massively Parallel Signature Sequencing*) para *Theileria annulata* (Bishop et al., 2005) e *Theileria parva* (Shah et al., 2006).

1.6.3 Transcriptoma de *Eimeria* spp.

A *Eimeria* apresenta um ciclo de vida complexo, envolvendo vários estágios de desenvolvimento, tais como, fases de proliferação que são críticas para a patogênese no hospedeiro e o ciclo sexual que é de suma importância para a transmissão, dispersão e geração da diversidade genética (Schmatz, 1997; Belli et al., 2005). Entretanto, pouco se sabe sobre os conjuntos de genes mais importantes e expressos nestes estágios do ciclo evolutivo.

Vários projetos para geração de ESTs de *Eimeria* de galinha doméstica abordando diferentes estágios do ciclo de vida destes parasitas foram realizados. Até o momento foram depositadas no *GenBank* aproximadamente 39.454 sequências ESTs (*Expressed Sequence Tags*), das quais cerca de 88,7% são de *E. tenella*, 6,9% de *E. maxima* e 3,6% são de *E. acervulina*. Dentre os estágios evolutivos estudados, os merozoítos e esporozoítos são os mais freqüentes, representando quase 75% do total de sequências ESTs depositadas até o momento.

Grande parte das sequências ESTs disponíveis são provenientes da parceria WashU-Merck estabelecida entre a *Washington University* e a Merck que produziu 27,500 ESTs de *E. tenella*. Cerca de 1.000 sequências foram produzidas pela *University Kebangsann Malaysia* (Wan et al., 1999; Ng et al., 2002) e 499 ESTs foram geradas a partir de oocistos esporulados e não esporulados pelo USDA (*United States Departamento of Agriculture*) (Miska et al., 2004). *Welcome Trust Sanger Institute* gerou mais de 8.000 ESTs de oocistos não esporulados, esporozoítos e merozoítos de primeira geração (dados não publicados).

Além destas ESTs convencionais, o nosso grupo, como parte integrante do Consórcio Internacional do Genoma de *E. tenella* foi responsável pelo sequenciamento de mais de 15.000 ESTs do tipo ORESTES (*Open Reading Frame EST*) para cada uma das três principais espécies de *Eimeria* de galinha doméstica: *E. acervulina*, *E. maxima* e *E. tenella*.

O primeiro estudo de expressão gênica em larga escala de *Eimeria* de galinha doméstica foi realizado na década de 90. Foram analisados cerca de 500 ESTs provenientes de merozoítos de segunda geração de *E. tenella*. Observou-se que 47,7% das proteínas eram similares a proteínas encontradas na base de dados, entre as quais, proteínas de ribossomo, quinases, oxidases, desidrogenases,

sintetases, entre outras, e um grupo remanescente de ESTs de proteínas putativas de funções diversas. Das sequências geradas, 14,3% eram semelhantes a genes previamente identificados em *E. tenella*. As sequências mais abundantes eram similares a proteínas envolvidas com o processo invasivo (Wan et al., 1999).

Estudo muito similar foi realizado pelo mesmo grupo, onde foram analisados 556 ESTs provenientes de esporozoítos de *E. tenella*. Apenas 27,3% das sequências obtidas apresentaram resultados de BLAST positivo, dentre estas, 22,5% eram genes nunca relatados em *E. tenella*. Após busca por similaridade por BLAST foi observado que esporozoítos e merozoítos apresentavam padrões distintos, denotando transcrição estágio específica. A maior parte dos transcritos de esporozoítos era associada ao crescimento celular, divisão celular e síntese de DNA, enquanto que, os transcritos de merozoítos estavam envolvidos com a expressão de genes e proteínas (Ng et al., 2002).

Padrão semelhante foi observado por Li et al. (2003), após a análise de mais de 55.000 ESTs de organismos do Filo Apicomplexa incluindo cerca de 13.500 ESTs provenientes de *Eimeria tenella* depositadas no banco de dados dbEST/GenBank. Estes autores observaram que 80% das sequências provenientes de esporozoítos e merozoítos de *Eimeria tenella* era estágio-específica. Similar proporção de transcritos estágio-específicos foi encontrada quando formas assexuais de *Plasmodium falciparum* foram analisadas (Li et al., 2003).

Um estudo mais recente também relatou que a expressão dos genes de esporozoítos e merozoítos de *E. tenella* é estágio específica. Após a geração de 35000 tags empregando SAGE foi observado que em merozoítos, grande parte dos produtos proteicos diferencialmente expressos estava relacionada à tradução, modificação e manutenção da conformação das proteínas e processos de ligação. Enquanto que em esporozoítos, os transcritos eram similares a histonas, proteínas associadas a transporte e atividade catalítica (Novaes, 2009).

Diferentemente do relatado em *E. tenella*, para *E. bovis* um estudo empregando *DD* (*Differential display*) demonstrou que grande parte dos transcritos de esporozoítos e merozoítos de primeira geração são compartilhados entre si (Abrahamsen et al., 1995).

Em *E. acervulina* o perfil transcricional de merozoítos de segunda geração foi também estudado. Após a montagem e análise de 1847 ESTs, foi observado que 29% das sequências geradas eram semelhantes às sequências presentes na bases

de dados de ESTS (dbEST). Dentre estas, a maioria era similar a proteínas hipotéticas, proteínas ribossômicas, proteínas envolvidas no metabolismo, replicação do DNA, transcritos, chaperonas e proteína de choque térmico (Miska et al., 2008).

Demais trabalhos avaliando a expressão de esporozoítos e esquizontes de primeira geração em *E. tenella* (Refega et al., 2003), bem como busca de possíveis candidatos vacinais em ESTs de esporozoítos e merozoítos de segunda geração desta mesma espécie (Klotz et al., 2005) foram também realizados. Entretanto há poucos trabalhos que estudam a expressão gênica em outras fases do ciclo de vida, como por exemplo, oocistos.

O primeiro artigo publicado sobre a expressão gênica em larga escala em oocistos de *Eimeria* foi realizado em 1993. Ao estudar *E. bovis*, os autores mostraram que o nível de transcrição de determinados genes varia ao longo da esporulação (Abrahamsen et al., 1993).

Em *E. tenella* também foi verificado que a transcrição de genes é diferente entre oocistos esporulados e não esporulados. Após a análise comparativa de ESTs de oocistos com os provenientes de esporozoítos e merozoítos depositados no *GenBank*, foi verificado que a metade dos transcritos de oocistos esporulados era também expressa em esporozoítos e merozoítos, enquanto que para oocistos não esporulados, 79% dos transcritos não foram detectados em outros estágios de desenvolvimento. Segundo os autores, não é surpresa que oocistos esporulados compartilhem semelhanças aos estágios invasivos, já que cada oocisto esporulado contém 8 esporozoítos no seu interior. Após análise de similaridade empregando o programa BLAST foi verificado que a maioria dos transcritos de oocistos esporulados era codificante para proteínas de micronema (Miska et al., 2004).

Proteínas de micronema, bem como proteína da família BT1 e proteínas ribossomais também foram observadas em transcritos obtidos de bibliotecas de cDNA de oocistos não esporulados, oocistos esporulados, esporozoítos e merozoítos (Wang et al., 2007).

O perfil de expressão gênica de oocistos não esporulados, oocistos esporulados e esporozoítos de *E. tenella* também foi estudado empregando hibridização subtrativa (SSH) e *microarray*. Foi observado que a maioria dos genes diferencialmente expressos era relacionada ao processo de esporulação, invasão

celular, interação parasita-hospedeiro, metabolismo celular, imunoregulação, tradução de sinal e regulação do ciclo celular (Han et al., 2010).

1.6.4 Transcriptoma de *Eimeria maxima*

Há poucos estudos sobre a expressão gênica nas diferentes fases do ciclo de vida de *E. maxima*.

O primeiro trabalho que avaliou a expressão gênica em larga escala em *E. maxima* foi realizado em 2006 por Basak e colaboradores. Estes autores empregaram DD (*Differential display*) com o objetivo de analisar o perfil transcricional de oocistos de duas cepas imunologicamente distintas, em diferentes fases da esporulação. Apesar de muitos transcritos estarem presentes em ambas as cepas e nos diferentes estágios estudados, foi verificado que um dos genes era transcrito somente numa das cepas independentemente da fase estudada (Basak et al., 2006).

Com o objetivo de realizar estudos comparativos entre cepas precoces e parentais de *E. maxima*, foi realizada uma análise da expressão gênica de oocistos esporulados de ambas as linhagens utilizando *Supression Subtractive Hybridization* (SSH). Um total de 360 ESTs foi obtido, destes, 21 genes eram mais transcritos na linhagem parental e 11 nas linhagens precoces. A maioria dos genes diferencialmente expressos era similar às proteínas relacionadas com invasão, mudança de estágio de desenvolvimento, esporulação, infectividade, virulência e resistência à drogas (Dong et al., *in press*).

Recentemente, o transcriptoma de *E. maxima* foi estudado a partir de um total de 3.168 ESTs de merozoítos que ao serem montados, resultaram em 416 *contigs* e 964 *singlets*. Após a anotação automática, 48,2% dos transcritos eram semelhantes às proteínas já descritas para organismos do Filo Apicomplexa, principalmente *Eimeria* spp. A maioria destas proteínas era relacionada à tradução, citoesqueleto, glicólise, proteínas de superfície, proteínas de sinalização, transporte, dobramento de proteína (*foldng*), metabolismo e proteínas de micronema. Estudos comparativos entre os transcriptomas das três espécies de *Eimeria* de galinha mais relevantes (*E. maxima*, *E. acervulina* e *E. tenella*) mostrou que há transcritos parcialmente, ou até

mesmo não compartilhados entre estas espécies. Dentre os transcritos mais abundantes, diferentes antígenos de superfície e uma proteína hipotética foram detectados nas três espécies de *Eimeria*. Ainda dentro deste grupo, duas proteínas (GAPDH 1 e antígeno de superfície MZ92/120) foram observadas em duas espécies de *Eimeria* (Schwarz et al., 2010).

Estudos comparativos do transcriptoma de *E. tenella*, *E. maxima* e *E. acervulina* foram também realizados pelo nosso grupo, a partir de sequências ORESTES provenientes de diversos estágios de desenvolvimento do ciclo destes parasitas. Dentro deste projeto, o presente trabalho foi desenvolvido com objetivo de se obter um perfil de expressão gênica qualitativo em larga escala de oocistos de *Eimeria maxima* em três fases da esporulação (oocistos não esporulados, parcialmente esporulados e esporulados).

O estudo de genes diferencialmente expressos nas fases da esporogonia poderá não somente contribuir para um melhor entendimento do processo de formação do esporozoítio e por conseqüente a invasão celular, como também no futuro poderá ser utilizado como alvo terapêutico, permitindo assim o desenvolvimento de novas estratégias de controle da doença.

2 OBJETIVOS

2.1 Objetivo principal

- Analisar comparativamente o perfil de transcrição gênica de diferentes fases da esporulação de *Eimeria maxima*.

2.2 Objetivos específicos

- Construção de bibliotecas de ESTs tipo ORESTES de oocistos não esporulados, oocistos parcialmente esporulados e oocistos esporulados de *Eimeria maxima*.
- Análise do perfil de transcrição gênica das diferentes fases da esporulação de *Eimeria maxima*
- Identificação dos transcritos diferencialmente expressos mais abundantes e seus padrões de transcrição.
- Validação dos dados obtidos por ORESTES para um pequeno conjunto de genes empregando-se a técnica de real-time PCR.

3 MATERIAIS E MÉTODOS

3.1 Preparo e tratamento das soluções

Todas as soluções foram preparadas utilizando-se água purificada por osmose reversa, deionizada e tratada com luz UV através do Sistema RiOs e Milli-Q Synthesis da Millipore (água Milli-Q). As soluções foram autoclavadas a 121 °C por 15 minutos para a eliminação de possíveis contaminações.

Para reagentes termossensíveis, utilizou-se água previamente autoclavada e as soluções preparadas foram esterilizadas por filtração (filtro 0,22 µm - Millipore).

Para a purificação de RNA mensageiro, todo o material utilizado foi adquirido ou previamente tratado de forma a não conter RNAses. Para isto, todas as soluções foram preparadas com água previamente tratada com dietilpirocarbonato de sódio (DEPC) 0,1% por 2 horas a 37 °C e posteriormente autoclavadas. Toda a vidraria utilizada foi incubada em forno Pasteur a 180 °C por 8 horas e o material plástico foi tratado com solução 0,1% DEPC por 2 horas a 37 °C e em seguida autoclavado. Bancadas e fluxos utilizados nas etapas de manipulação de RNA, foram limpos com solução de dodecil sulfato de sódio (SDS) 0,5%.

3.1.1 Soluções

- Tampão de amostra tipo II, 6X concentrado - 15% (p/v) Ficoll (tipo 400) em H₂O; 0,25% (p/v) azul de bromofenol; 0,25% (p/v) xileno ciano.
- Tampão de extração - Tris-HCl 10 mM pH 8,0; EDTA 50 mM pH 8,0.
- Tampão de sequenciamento 2,5X concentrado - Tris-HCl 200 mM pH 9,0; MgCl₂ 5,0 mM.
- TBE 10X concentrado - Tris-borato 900 mM; EDTA 20 mM em de H₂ O, pH 8,0.
- TE - Tris-HCl 10 mM pH 8,0; EDTA 1 mM pH 8,0.
- Tampão MOPS - 0,2 M de ácido 3-(N-morfolino propano sulfônico, 20 mM acetato de Sódio; 10mM EDTA, pH 8,0)

3.1.2 Meios de cultura

- Ágar LB – Caldo LB com 15 g/L de ágar.
- Caldo LB - 10 g/L de Bacto-triptona; 5 g/L de extrato de levedura e 10 g/L de NaCl, pH 7,0.
- Caldo 2YT - 16 g/L de Bacto-triptona; 10 g/L de extrato de levedura e 5 g/L de NaCl, pH 7,0.
- GYT- 0,125% (p/v) de extrato de levedura; 0,25% (p/v) de triptona; 10% (v/v) glicerol.
- SOB – 20 g/L de bacto-triptona; 5g/L de extrato de levedura; 0,5 g de NaCl; 2,5 mM de KCl; 10 mM de MgSO₄, pH 7,0.
- SOC – 20 g/L de bacto-triptona; 5g/L de extrato de levedura; 0,5 g de NaCl; 2,5 mM de KCl; 10 mM de MgSO₄, 20 mM de glicose, pH 7,0.

3.2 Condições de manutenção das aves

Foram utilizados machos de postura da linhagem Bovans White (Hendrix Poultry Breeders) cedidos pela Granja Kunitomo (Mogi das Cruzes, SP) com um dia de idade.

As aves foram alojadas no biotério do Laboratório de Biologia Molecular de Coccídias do Departamento de Parasitologia, Instituto de Ciências Biomédicas, USP, em gaiola do tipo bateria contendo bebedouro automático e aquecedor elétrico. Os animais foram mantidas com água e ração livre de anticoccidianos *ad libitum* em ambiente com ventilação constante e ar filtrado na entrada e saída das salas. Após 3 a 6 semanas de idade, as aves foram transferidas para uma sala de infecção, onde foram infectadas e alojadas em gaiolas de 50 x 50 x 50 cm.

Considerando-se que várias espécies de *Eimeria* podem infectar simultaneamente um mesmo hospedeiro, e como no nosso laboratório, algumas das espécies são manipuladas e propagadas com frequência, as condições de criação e manutenção das aves devem ser controladas. Desta forma, a sala de criação tem pressão de ar positiva enquanto que as salas de infecção têm pressão de ar

negativa, além disto, as salas de infecção possuem uma rede de gás amônia para esterilização do ambiente e eliminação dos oocistos residuais. Além destas medidas, durante o período de criação, as fezes das aves foram monitoradas quinzenalmente quanto à presença de oocistos de *Eimeria* que acometem a galinha doméstica.

3.3 Propagação dos parasitas

3.3.1 Infecção das aves com oocistos de *Eimeria maxima*

Para a propagação dos parasitas, as aves foram infectadas com oocistos esporulados de *Eimeria maxima* cepa H (Houghton, Inglaterra) por via oral com auxílio de uma seringa descartável conectada a uma sonda de polietileno, na dose de 5×10^3 oocistos/ave (Long et al., 1976).

A cepa de referência H, utilizada neste estudo, foi gentilmente cedida pelo Dr. Martin W. Shirley (*Institute for Animal Health*, Reino Unido).

3.3.2 Purificação dos oocistos

Para purificação de oocistos de *Eimeria maxima*, cepa H, foi utilizada uma variação do método de Long et al. (1976). As fezes das aves foram coletadas durante o período de patência, 120 a 196 horas, após a infecção.

Após a coleta, as fezes foram homogeneizadas com um volume de água destilada, utilizando-se um agitador mecânico com hélice (agitador naval). Em seguida, as fezes foram peneiradas utilizando-se uma tela de aço inoxidável com malha de 40 *mesh* (poros de 425 μm) e lavadas por 3 a 4 vezes com água destilada.

O material purificado foi centrifugado a 3500 *g* por 5 minutos (centrífuga Sorvall, modelo Super T-21), o sobrenadante foi descartado e o sedimento ressuspenso em solução salina saturada. A suspensão foi submetida à centrifugação a 2.500 *g* por 5 minutos com desaceleração 133g cm/min^2 . A fase

superior, similar a uma nata e constituída de oocistos, foi cuidadosamente removida com uma seringa de vidro acoplada a uma agulha 40x12 mm e transferida para um novo tubo. Este processo foi repetido mais duas vezes. As fases contendo os oocistos foram diluídas em água destilada e centrifugadas a 2500 g por 5 minutos com desaceleração $831g \text{ cm/min}^2$, este último passo foi repetido por mais três vezes, até a remoção total da solução salina saturada.

3.3.3 Esporulação dos oocistos

Os oocistos de *Eimeria maxima* foram ressuspensos em solução 2% (p/v) de dicromato de potássio em uma concentração final de $2,5 \times 10^5$ oocistos/mL (Shirley, 1995). A esporulação dos oocistos foi realizada em uma câmara de incubação B.O.D. (modelo 347 CD, Fanem, São Paulo, Brasil). Os oocistos foram mantidos a 28 °C sob oxigenação forçada através borbulhamento de ar por 72 horas (Long et al., 1976) Em seguida, o volume da suspensão foi reduzido e os oocistos foram estocados a 4 °C.

Para a obtenção dos oocistos na fase intermediária da esporulação foi realizada uma cinética de esporulação, visando identificar o período no qual era possível obter-se um número máximo de oocistos na fase de esporoblasto. Para tanto, durante um período de 72 horas, alíquotas da suspensão de oocistos foram retiradas periodicamente para análise morfológica e contagem da proporção de oocistos nas diferentes fases da esporulação em câmara Fucs-Rosenthal visualizada em microscópio Nikon (modelo Eclipse E800).

3.4 Tratamento dos oocistos com hipoclorito de sódio

Visando a remoção de impurezas e bactérias contaminantes que poderiam prejudicar as etapas de extração do DNA e RNA mensageiro de *Eimeria maxima*, as amostras contendo os oocistos foram tratadas com hipoclorito de sódio. Inicialmente a suspensão de oocistos foi centrifugada a 2.500 g por 5 minutos com

desaceleração $831g \text{ cm/min}^2$ (Sorvall, modelo Super T-21). O sobrenadante foi dispensado e o sedimento foi lavado de duas a três vezes com água destilada, até a remoção total do dicromato de potássio. O sedimento foi ressuspensão em hipoclorito de sódio 10-12% (cloro ativo) e incubado por 10 minutos a 4°C . Em seguida, a amostra foi diluída, por no mínimo, 10 vezes em água destilada, centrifugada a $3.500 g$ por 5 minutos com desaceleração $831g \text{ cm/min}^2$. O sobrenadante foi dispensado e o sedimento foi ressuspensão em água destilada. Esta última etapa foi repetida por no mínimo três vezes até a retirada total do hipoclorito.

3.5 Verificação de contaminação inter-específica

Além da realização dos exames parasitológicos periódicos das fezes das aves mantidas no nosso biotério, todas as amostras de oocistos purificados foram submetidas a um ensaio de PCR multiplex capaz de diagnosticar simultaneamente as sete espécies de *Eimeria* que infectam a galinha doméstica. Este teste baseia-se na utilização de marcadores SCAR (*Sequence Characterized Amplified Regions*) espécie-específicos (Fernandez et al., 2003b) e foi empregado com o objetivo de monitorar uma possível contaminação inter-específica no material purificado, já que o nosso grupo trabalha também com outras espécies de *Eimeria* que infectam a galinha doméstica. .

Para tanto, o DNA foi extraído a partir de 5×10^6 oocistos previamente tratados com hipoclorito de sódio conforme descrito no item 3.4, e ressuspensos em tampão de extração. Os oocistos foram ressuspensos em $300 \mu\text{L}$ tampão de extração (Tris-Cl 10 mM pH 8.0; EDTA 50 mM pH 8.0) e rompidos totalmente através de agitação em vortex (Daigger Vortex Genie 2™, A. Daigger & Co, Inc., EUA) com $\frac{1}{2}$ volume de pérolas de vidro ($0,4 \text{ mm}$ de diâmetro, Sigma-Aldrich Corp., EUA). Após centrifugação a $20.000 g$ (centrífuga Eppendorf, modelo 5417R, Hamburg, Alemanha) durante 10 minutos a 4°C , o sobrenadante foi recolhido e tratado com RNase A ($20 \mu\text{g/mL}$) a 37°C por 1 hora. A seguir foram adicionados SDS ($0,5\%$ p/v) e Proteinase K ($100 \mu\text{g/mL}$) e a amostra foi incubada a 50°C durante 2 horas. O DNA foi purificado pelo método fenol/clorofórmio (Sambrook e Russell, 2001).

Os ensaios de PCR multiplex foram realizados de acordo com Fernandez et al. (2003). Para todas as reações foram realizados controles positivo e negativo. Os produtos amplificados foram submetidos à eletroforese em gel de agarose 1,5%.

Para a verificação dos resultados em gel, após a amplificação dos produtos, o tampão de amostra do tipo II, 6X concentrado foi adicionado às amostras. Em seguida as amostras foram aquecidas a 65 °C por 5 minutos para desnaturação do DNA amplificado e aplicadas em gel de agarose. A corrida foi realizada em TBE 1X contendo brometo de etídeo na concentração de 0,5 µg/mL sob uma tensão constante (10 v). O gel foi analisado sob luz UV e para aquisição das imagens, utilizou-se um sistema de foto documentação digital ChemImage 4400 (Alpha Innotech Corp., Santa Clara, Califórnia, EUA).

3.6 Extração de RNA mensageiro de oocistos de *Eimeria maxima*

O RNA mensageiro (RNAm) dos oocistos não esporulados, parcialmente esporulados (fase de esporoblasto) e esporulados foram extraídos utilizando-se o kit µMACS mRNA isolation (Miltenyi Biotec Inc, Auburn, California, EUA), conforme instruções do fabricante. Este kit baseia-se no uso de oligo dT ligado a partículas paramagnéticas, permitindo o isolamento do RNAm a partir de células em uma única etapa.

Para a extração do RNAm de *Eimeria maxima*, os oocistos dos diferentes estágios da esporulação foram lavados em H₂O DEPC, centrifugados a 12.000g por 5 minutos e o sobrenadante foi descartado. Os oocistos foram ressuspensos em 500 µL tampão de extração RNase free e rompidos por meio de agitação em vortex (Daigger Vortex Genie 2) com ½ volume de pérolas de vidro (0,4 mm de diâmetro), por no máximo 6 minutos. A fase de quebra foi monitorada por microscopia óptica. Após a quebra, a amostra foi centrifugada a 20.000 g por 10 minutos (centrífuga Eppendorf), o sobrenadante foi recolhido, transferido para um novo tubo contendo 500 µL *Lysis buffer* (kit µMACS mRNA isolation) procedendo-se a seguir a extração do RNAm conforme instrução do fabricante.

Para avaliação quantitativa e qualitativa do RNAm, um décimo do material purificado foi submetido à eletroforese em gel de agarose 1% em tampão 1x MOPS, na ausência de brometo de etídeo. Após a corrida, o gel foi corado com brometo de etídeo (0,5 µg/mL) em MOPS 1x por 10 minutos e visualizado sob luz UV.

Para eliminação de possíveis contaminações com DNA genômico, as amostras de RNA mensageiro purificado, foram submetidas a tratamento com DNase. Para o ensaio utilizou-se 1U/10µL de reação de DNase RQ1 livre de RNase (Promega Corporation, Madison, Wisconsin, EUA), na presença de tampão de DNase 1x concentrado, DTT (2 mM final) e inibidor de RNase, RNasin (Promega) (0,5 U/µL de reação). A amostra foi incubada a 37 °C por 30 minutos e a reação de digestão foi interrompida pela adição de *Stop Solution* e incubação a 65 °C por 5 minutos.

A quantificação do RNA mensageiro foi realizada em espectrofotômetro (GeneQuant-pro™, GE Healthcare Biosciences, Pittsburgh, EUA), medindo-se a absorbância das amostras em 260/280nm. Para a realização de ensaios de real time PCR (item 3.19) o RNA mensageiro foi quantificado no fluorômetro Qubit™ (Invitrogen) cujo fluoróforo (Qubit™ RNA Assay Kit, Invitrogen) somente se torna fluorescente quando se liga especificamente à moléculas de RNA. Este tipo de medição é mais acurado que o anterior já que a quantificação reporta somente a concentração da molécula de interesse. Em seguida, o RNA mensageiro foi aliquoteado e estocado a -80 °C.

3.6.1 Verificação de contaminação do RNA mensageiro com DNA genômico

Além do tratamento com DNase (ver item 3.6), a presença de DNA genômico contaminante nas amostras de RNA mensageiro foi descartada pela realização de ensaio de PCR empregando *primers* baseados em marcadores SCAR espécie-específicos para *E. maxima* (Fernandez et al., 2003b). A amplificação desta região gera um fragmento não codificante de 272 pb.

Os testes de PCR foram realizados em um volume final de 25 µL contendo 200 µM de dNTP, 1x de tampão, 1 U de BIOLASE™ DNA polimerase (Bioline, Londres, Reino Unido), 1,5 mM de MgCl₂, 1 µM de cada *primer* específico para *E.*

maxima (Forward (F) - GGG TAA CGC CAA CTG CCG GGT ATG; Reverse (R) – AGC AAA CCG TAA AGG CCG AAG TCC TAG A), empregando as condições de ciclagem segundo Fernandez et al. (2003). Para todas as reações foram realizados controles positivo e negativo. Os produtos amplificados foram submetidos à eletroforese em gel de agarose 1,5%.

3.6.2 Verificação da integridade do RNA mensageiro

Com o objetivo de verificar a integridade do RNA mensageiro purificado, foi realizada uma reação de transcrição reversa (RT-PCR) empregando *primers* específicos para uma região do gene de proteína TFP-250 (Genbank: AY239227), membro da família das proteínas TRAP, que está presente em todos os estágios assexuais de *Eimeria maxima* (Smith et al., 1994; Witcombe et al., 2003). A amplificação desta região gera um fragmento de 900 pb. A síntese de cDNA foi realizada a partir de *primer* específico, utilizando-se a enzima SuperScript™ II *Reverse Transcriptase* (Invitrogen, Life Technologies, Carlsba, California, EUA), conforme instruções do fabricante. A amplificação do cDNA foi realizada utilizando-se 10% do volume da reação de síntese de cDNA, 1 U de Platinum *Taq* DNA Polymerase (Invitrogen), 1x tampão da enzima, 1,5 mM de cloreto de magnésio, 100 µM de dNTP e 1 µM de cada primer (Forward (F) – GCT GCT CTG TAA ACG CCA CTT G; Reverse (R) – CAC TCG TCG ATG TCG GTA CAC C). As condições da reação consistiram em uma desnaturação inicial de 95 °C por 5 minutos; 30 ciclos de 95 °C por 1 minuto, 57 °C por 1 minuto e 72 °C por 1 minuto e uma etapa final de extensão a 72 °C por 5 minutos. Para verificação da amplificação, um terço de cada amostra amplificada foi submetido à eletroforese em gel de agarose 1,5%.

3.7 Construção de bibliotecas de cDNA do tipo ORESTES

3.7.1 Obtenção e manipulação dos perfis

Para a obtenção dos perfis de cDNA foi empregada a técnica denominada ORESTES (ORF ESTs). Este método baseia-se em PCR de baixa estrincência utilizando *primers* arbitrários para amplificação dos cDNA, há maior cobertura da região central da sequência dos RNAs mensageiros, a qual contém a maior parte da porção codificadora dos genes (Dias Neto et al., 2000).

Para a síntese de cDNA de cada estágio evolutivo foram usados 10ng de RNA mensageiro. A reação de transcrição reversa foi realizada a 37 °C por 90 minutos em tubos com volume final de 3,5 µL contendo 60 U de enzima de transcrição reversa SuperscriptTM II (Invitrogen), 250 µM de dNTP, 20 mM de DTT e 22,5 pmol de primers arbitrários de 17 a 31 nucleotídeos.

O cDNA obtido foi amplificado empregando o *kit PuReTaq Ready-To-GoTM PCR Beads* (GE). Os *beads* foram dissolvidos em 25 µL de H₂O MilliQ autoclavada e em seguida foram acrescidos 3,5µL de cDNA.

As condições da reação consistiram em PCR *touch down* de 45 ciclos (Verjovski-Almeida et al., 2003). Após desnaturação inicial a 75 °C por 5 minutos, procedeu-se dois ciclos de 94 °C por 30 segundos, 52°C por 1 minuto e 72 °C por 1 minuto. A ciclagem foi repetida com uma progressiva redução da temperatura de anelamento, um grau a cada dois ciclos, até atingir 45 °C. Em seguida, foram realizados 26 ciclos de 94 °C por 30 segundos, 48 °C por 1 minuto e 72 °C por 1 minuto, com uma extensão final de 72 °C por 7 minutos.

Para verificação da amplificação e seleção do tamanho dos fragmentos, os perfis foram submetidos à eletroforese em gel de agarose 1% corados com brometo de etídeo. Os perfis de DNA foram visualizados sob lâmpada UV portátil (366 nm) e os fragmentos de tamanho entre 400 e 1000 pb foram selecionados e excisados do gel. A purificação do DNA foi realizada utilizando-se o *kit GFXTM PCR DNA and Gel Band Purification Kit* (GE), segundo instruções do fabricante.

3.7.2 Clonagem dos perfis

Após a purificação dos perfis de DNA, os fragmentos foram ligados ao vetor *pGEM – T Easy* (Promega), utilizando 35 ng do vetor, 50 ng de amostra e 3 U de T4 DNA Ligase. A reação foi incubada a 16 °C por 16 horas, precipitada e o *pellet* foi ressuspensão em H₂O MilliQ.

As bibliotecas de cDNA foram clonadas em bactérias *E. coli*, eletrocompetentes, cepa DH10B, utilizando o vetor *pGEM – T Easy* (Promega). Para tanto, o vetor foi inserido nas bactérias por eletroporação utilizando o eletroporador Gene Pulser II™ (Bio-Rad), calibrado para um pulso de 1.8 KV, capacitância de 50 µF e resistência de 100 Ω. Após a transformação, um total de 300µL de meio SOC foi adicionado e o material foi incubado a 37 °C por 1 hora. Em seguida, as bactérias foram plaqueadas em placas de Petri contendo meio LB ágar com ampicilina (100 µg/mL final), 90 µL de X-Gal (2% p/v em dimetilformamida) e 10 µL de IPTG (0,1 M). As placas foram incubadas a 37 °C por 14 a 16 horas.

Após a incubação, as colônias isoladas de coloração branca (que continham o inserto) foram picadas com palitos estéreis e inoculadas nos orifícios de microplacas de 96 cavidades contendo 100 µL de caldo TB adicionado de ampicilina (100 µg/mL final). Em seguida, as bactérias foram incubadas a 37 °C por 14 horas e, para estocagem das bactérias, foram adicionados em cada poço um volume de glicerol 30% (v/v) estéril. As microplacas foram estocadas a –80 °C até o momento do uso.

3.7.3 PCR de colônia

Para obtenção dos DNA moldes os insertos presentes nas bactérias transformadas foram amplificados por PCR de colônia. Esta técnica se baseia no princípio de utilização de uma temperatura inicial de desnaturação do DNA alta o suficiente para promover a lise bacteriana e liberação do DNA.

As reações de amplificação foram realizadas em um volume final de 15 µL, utilizando-se 1 µL de cada suspensão bacteriana transferida das microplacas de cultura com o auxílio de um replicador de 96 agulhas. Para cada reação de PCR

foram utilizados 2,3 μM de cada primer M13 *forward* (TGTAACGACGGCCAGT) e M13 *reverse* (CAGGAAACAGCTATGACC), 125 μM de dNTP, 1 U de enzima BIOLASE™ DNA polimerase (Bioline) e tampão da enzima 1x. As condições de amplificação consistiram em uma desnaturação inicial de 95 °C por 4 minutos; 40 ciclos de 95 °C por 45 segundos, 55 °C por 45 segundos e 72 °C por 1 minuto; e uma etapa final de extensão a 72 °C por 5 minutos. Para verificação da amplificação, um quinto de cada reação de PCR foi submetido à eletroforese em gel de agarose 1,2%. As microplacas foram estocadas a -80 °C até o momento do uso.

3.8 Reação de sequenciamento

A reação de sequenciamento foi realizada pelo método de terminação de cadeia de Sanger (Sanger et al., 1977), utilizando-se o *ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit version 3.1* (Applied Biosystems by Life Technology, Foster City, Califórnia, EUA).

Para tanto, o DNA resultante do PCR de colônia foi diluído 1:10 em água e o sequenciamento foi realizado em microplacas de 96 poços utilizando-se *primers* que flanqueiam os insertos presentes no vetor *pGEM-T Easy* (descritos no item 3.7.2).

As reações de sequenciamento foram realizadas em um volume final de 20 μL , utilizando-se 1 μL do PCR de colônia com o auxílio de um replicador de 96 agulhas. Para cada reação de sequenciamento foram utilizados 30 μM de primer M13 (5' TGTAACGACGGCCAGT 3') ou M13R (5' CAGGAAACAGCTATGACC 3'), 1 μL de *ABI BigDye Terminator Cycle Sequencing Ready Reaction Kit version 3.1* (Applied Biosystems) e tampão da enzima 1x. A condição de reação empregada foi de 40 ciclos de 96 °C por 10 segundos, 52 °C por 20 segundos e 60 °C por 4 minutos.

Após a reação, os terminadores não incorporados foram removidos por precipitação diferencial conforme instruções do fabricante (PE Applied Biosystems, 1998). Após a precipitação o DNA foi resolubilizado em 10 μL de formamida 10% (v/v).

As amostras foram desnaturadas a 95 °C por 5 minutos e aplicadas em um sequenciador automático de DNA modelo ABI 3700 (Applied Biosystems) de 96 capilares.

3.9 Pré-processamento das sequências

Todas as sequências geradas foram submetidas a um processamento seriado de múltiplas etapas (“*pipeline*”) na plataforma EGene (Durham et al., 2005) desenvolvido pelo nosso grupo de bioinformática e que constituiu das seguintes etapas:

- **Atribuição de bases (“base calling”) e avaliação da qualidade.** Para essa etapa utilizou-se o componente `upload_traces.pl` do EGene, acoplado ao programa Phred (Ewing et al., 1998; Ewing et al., 1999).

- **Mascaramento das sequências dos *primers* de ORESTES:** cada biblioteca foi construída utilizando diversos *primers* arbitrários. As sequências dos *primers* foram mascaradas utilizando-se o componente `mask_cross_match.pl` do EGene e o programa Cross_match (Phil Green, não publicado).

- **Mascaramento de bases de vetor:** todas as sequências nucleotídicas foram mascaradas contra a sequência do vetor *pGEM T-Easy* (Promega). Utilizou-se o componente do EGene `mask_cross_match.pl` acoplado ao programa Cross_match do pacote Phred/Phrap/Consed.

- **Filtragem por qualidade.** Nesta etapa procurou-se eliminar todas as sequências que não atendiam o critério de qualidade Phred. Para ser aceita, uma leitura deveria apresentar pelo menos uma sequência contínua de 100 pb com 85% das bases com qualidade Phred igual ou maior do que 15. Para isso, utilizou-se uma abordagem de janela deslizante. O programa utilizado para essa etapa foi o `filter_quality.pl` do sistema EGene.

- **Aparamento das pontas.** Após a filtragem por qualidade, as sequências das pontas tiveram suas bases de qualidade ruim aparadas. Para isto, usou-se o componente `trimming.pl` do EGene. As regiões da sequência para serem

aceitas, tinham que apresentar dentro uma janela deslizante de 30bp, 80% das bases com índice Phred igual ou maior do que 10,

- **Filtragem por tamanho.** Após o processamento e aparamento de pontas, todas as sequências foram submetidas ao programa `filter_size.pl` do EGene. Somente foram aceitas sequências com mais de 70 pb.

- **Filtragem de sequências mitocondriais:** foi feita a filtragem de todas as sequências empregando o componente `filter_cross_match.pl`. Todas as sequências foram comparadas aos genomas mitocondriais de espécies de *Eimeria* determinadas previamente pelo laboratório (Romano, 2004). Como critério de positividade, a sequência tinha que apresentar os parâmetros de *Cross_match* “-minmatch 35 -penalty -1 -minscore 30”.

- **Filtragem de sequências de plastídeo:** todas as sequências foram filtradas utilizando o componente `filter_cross_match.pl`, que utiliza o programa *Cross_Match* para fazer buscas de similaridade. Como base de dados, foi usada a sequência nucleotídica do genoma de Apicoplasto de *E. tenella* (código de acesso AY217738). Foram consideradas positivas as sequências que apresentaram alinhamento utilizando-se os mesmos parâmetros de filtragem de sequências mitocondriais.

- **Filtragem de sequências ribossômicas:** foi feita a filtragem de todas as sequências com o componente `filter_blast.pl`, que é acoplado ao programa BLAST. Como base de dados, foi usada uma combinação de sequências ribossômicas e organismos do Superfilo Alveolata. Para uma sequência ser considerada positiva, tinha que atender três critérios: apresentar bloco de alinhamento com *e-value* de valor máximo de 10^{-20} , tamanho de bloco de no mínimo 90pb e porcentagem de similaridade de no mínimo 88% dentro do bloco.

- **Filtragem de sequências repetitivas:** nesta etapa também se utilizou o componente `filter_cross_match.pl`, e os mesmos critérios de estrigência adotados no item filtragem de sequências mitocondriais. Como base de dados, empregou-se o Repbase (<http://www.girinst.org/rebase/index.html>), uma base ampla de elementos repetitivos eucariotos.

- **Filtragem de sequências bacterianas:** utilizou-se o componente `filter_blast.pl`, e os mesmos critérios de estrigência adotados para a filtragem de sequências ribossômicas. Como base de sequências, utilizou-se os

genomas de *Escherichia coli* (NC_000913.1), *Xanthomonas campestris* pv. *campestris* (NC_003902.1), *Xanthomonas axonopodis* pv. *citri* (NC_003919.1) e *Xylella fastidiosa* 9a5c (NC_002488.1). A *E. coli* foi usada pelo fato de ter sido empregada em todas as etapas de clonagem e propagação dos clones recombinantes. As demais bactérias foram incluídas pelo fato de seu DNA ter sido amplamente manipulado no laboratório durante os diferentes projetos de sequenciamento de genomas do qual nosso grupo fez parte.

- **Filtragem de sequências de galinha doméstica:** utilizou-se o componente `filter_blast.pl` e os mesmos critérios de estringência adotados para a filtragem de sequências ribossômicas, com exceção de um valor de *e-value* 10^{-25} mais rigoroso. Como bases de dados utilizou-se as sequências do genoma de *Gallus gallus*.

- **Filtragem de sequências humanas:** utilizou-se o componente `filter_blast.pl` e os mesmos critérios de estringência para filtragem de sequências de galinha doméstica foram adotados. Como base se sequências utilizou-se as sequências do genoma humano montado, disponível em <http://genome.uscs.edu> ("*Golden Path*"). Esta etapa de filtragem contra o genoma humano foi adotada pelo fato desta técnica ser baseada em amplificação arbitrária de DNA e portanto, poderia amplificar o DNA dos manipuladores.

3.10 Agrupamento das sequências geradas por ORESTES

O agrupamento das sequências (*clustering*) foi realizado utilizando o programa CAP3 (Huang e Madan, 1999), empregando o parâmetro *p* (limiar de porcentagem de identidade de sobreposição) igual a 90 e *y* (faixa de cortes de pontas) igual a 250 (Ferro, 2008).

3.11 Verificação da posição relativa de transcritos ORESTES em cDNAs completos de *E. maxima*

Com o objetivo de avaliar a posição relativa dos transcritos ORESTES nas sequências de *E. maxima*, foram realizadas buscas de sequências gênicas completas (*full lenght*) no banco de dados GenBank, seguida por análise de similaridade (BLASTN) entre os transcritos ORESTES e os genes encontrados. Somente foram analisados os transcritos que apresentaram alta similaridade (e *value 0.0*). Durante a análise foi verificado se o transcrito mapeava nas regiões 5', central ou 3' do gene correspondente. Para averiguar os transcritos mapeavam na região codificante, foi realizado busca de ORFs nas sequências completas empregando o ORF Finder (*Open Reading Frame Finder* - <http://www.ncbi.nlm.nih.gov/projects/gorf/>)

3.12 Mapeamento dos *contigs* contra o genoma de *E. maxima*

O conjunto de *contigs* obtidos a partir das bibliotecas de ORESTES foi submetido a um mapeamento global contra a montagem do genoma de *E. maxima* cepa H, disponível no banco de dados EmaxDB: *The Eimeria maxima database* (<http://www.genomemalaysia.gov.my/emaxdb/>). Para tanto, análises de similaridade entre as sequências ORESTES e os *contigs* genômicos foram realizadas empregando BLASTN.

3.13 Anotação automática das sequências geradas por ORESTES

A anotação automática foi realizada pelo nosso grupo durante o trabalho de dissertação de mestrado da aluna Milene Ferro, sob orientação do professor Arthur Gruber (Ferro, 2008). Foi construído um *pipeline* de anotação automática que consiste das seguintes etapas.

- **Determinação das fases de leituras abertas (ORFs) e tradução proteica:** para esta etapa utilizou-se o componente `annotation_orf.pl`. Foram escolhidas as fases de leitura abertas contendo pelo menos 50 resíduos de aminoácidos, sem requerimento de códon de iniciação. Todas as ORFs foram traduzidas conceitualmente utilizando-se o código genético universal e sequências nucleotídicas e proteicas correspondentes armazenadas em relatórios.

- **Busca de similaridade:** para as buscas de similaridade foi utilizada um componente `annotation_blast.pl` acoplado ao programa BLAST (Altschul et al., 1997). Como base de dados foi utilizada a base nr (não redundante) do *GenBank*. Os quatro melhores alinhamentos com *e-values* inferiores a $1e-06$ e porcentagem de similaridade maior ou igual a 55% foram anotados.

- **Busca de domínios conservados:** as sequências proteicas geradas pelo componente `annotation_orf.pl` foram submetidas a uma busca de domínios conservados contra a base de dados CDD (Marchler-Bauer et al., 2007). Para isso utilizou-se o componente `annotation_rpsblast.pl` acoplado ao programa RPS-BLAST (Marchler-Bauer et al., 2002). Os alinhamentos com *e-values* inferiores a 10^{-6} foram anotados.

- **Busca de domínios proteicos:** as sequências proteicas foram submetidas a uma busca de motivos proteicos utilizando-se o componente do EGene `annotation_interpro.pl` acoplado ao programa Interpro-Scan (Mulder e Apweiler, 2007).

- **Busca de sequências de peptídeo-sinal:** para essa busca foram utilizados os componentes `annotation_signalP.pl` e `annotation_phobius.pl`, acoplados respectivamente aos programas SignalP (Bendtsen et al., 2004) e Phobius (Kall et al., 2004).

- **Busca de domínios transmembranares:** nessa busca foram utilizados os componentes `annotation_tmhmm.pl` e `annotation_phobius.pl`, acoplados respectivamente aos programas TMHMM (Krogh et al., 2001) e Phobius (Kall et al., 2004).

- **Busca de sequências de sítos de ancoras GPI:** nesta etapa utilizou-se o componente `annotation_dgpi.pl` acoplado ao programa DGPI (Kronegg e Buloz, 1999).

- **Seleção das ORFs anotadas:** foi utilizado o componente `report_ORF.pl` para selecionar ORFs com maior número de evidências tais como resultados de similaridade por BLAST com melhor *e-value*, presença de domínios conservados detectados pelo RPS- BLAST e motivos proteicos encontrados pelo InterproScan. Em caso de empate do número de evidências, o componente seleciona a ORF de maior tamanho.

- **Mapeamento dos termos GO (Gene Ontology):** uma vez concluído o pipeline de anotação automática, foi usado o componente do EGene `mapping_GO.pl` para realizar o mapeamento e quantificação dos termos GO nas três ontologias para cada sequência anotada.

3.14 Northern Digital de sequências ORESTES de *E. maxima*

Os cDNAs reconstruídos oriundos da montagem realizada pelo programa CAP3 foram utilizados para gerar perfis de expressão empregando *Northern* digital (Audic e Claverie, 1997). Considerando que cada cDNA reconstruído pode estar representado por sequências de diferentes estágios evolutivos, a abundância destas poderia refletir o nível relativo de transcrição destes cDNAs. Baseando-se nesta premissa, para cada cDNA reconstruído, as leituras foram contadas e computadas de acordo com o estágio evolutivo empregando um *script* em Perl desenvolvido pelo laboratório. Para a identificação dos genes diferencialmente expressos os dados foram submetidos à análise estatística utilizando o programa Kemp (item 3.16).

3.15 Agrupamento das sequências ORESTES utilizando o programa Simcluster

Com o objetivo de agrupar os *contigs* de diferentes estágios evolutivos de *E. maxima*, os dados oriundos das bibliotecas de ORESTES foram submetidos a um agrupamento empregando o programa Simcluster (*Clustering Software for SIMPLEX data*) (Vencio et al., 2007). Para a análise ser mais abrangente, além dos oocistos (Ou, Op e Os), *contigs* originários de merozoítos (Mz) também foram empregados.

As bibliotecas de Mz não fizeram parte deste trabalho e foram geradas pelo grupo dentro do projeto ORESTES de *E maxima*.

Para esta análise foi realizado um agrupamento hierárquico utilizando máxima distância de pareamento (*Maximum pairwise distance / linkage* completo) entre as bibliotecas empregando diferentes métodos de distância métrica (Euclidiano, Aitchison, Pearson *Pseudo-Distance* e *Uncentered Pearson Pseudo-Distance*). Os dendrogramas foram editados utilizando o programa FigTree (A. Rambaut, não publicado, disponível no sítio <http://tree.bio.ed.ac.uk/software/figtree>).

3.16 Análise estatística

Após o *Northern* digital os dados foram submetidos a uma ferramenta de análise estatística denominada Kemp (Varuzza, 2008). A análise é baseada em um teste frequentista desenvolvido para a comparação de perfis de expressão digital.

Para classificar se um *contig* é expresso diferentemente em um conjunto de bibliotecas empregou-se o *p-value* (p) que representa uma medida numérica da coerência dos dados observados em relação à hipótese nula (Cox, 1975; Kempthorn, 1976). Neste teste os valores de p foram calculados considerando a hipótese nula como a não diferenciação dos *contigs*. Além do cálculo dos valores de p , o Kemp calcula automaticamente o nível crítico (α) para cada *contig* em particular. Este valor é utilizado para se comparar os valores de p , e assim, aceitar ou rejeitar a hipótese nula. A escolha do nível crítico pode gerar dois tipos de erro:

- **Erro do tipo I:** Rejeitar a hipótese nula quando ela é verdadeira. Este erro também é chamado de falso positivo;
- **Erro do tipo II:** Aceitar a hipótese nula quando ela é falsa. Também chamado de falso negativo.

Para tanto, o programa utiliza um procedimento de decisão teórica que minimiza os erros do tipo I (α') e do tipo II (β) da função $a\alpha' + b\beta$. Neste caso foi escolhido $a=4$ e $b=1$, empregando assim uma maior estringência para escolha dos *contigs* diferencialmente expressos, minimizando o número de resultados falso positivos, pois o valor de α' , que considera a hipótese nula verdadeira (*contigs* não

diferencialmente transcritos) possui um peso maior do que o valor de β , onde a hipótese nula é considerada falsa (*contigs* diferencialmente transcritos).

O programa Kemp gera um arquivo do tipo csv (*comma-separated values*) no qual cada linha contém informações relativas a cada *contig* tais como: identificador; número de leituras de cada estágio evolutivo (Ou, Op e Os); poder de expressão (PE), ou seja, soma do número de leituras; valores de p (valor de significância); α (nível crítico); *score* da sequência $\{10 \times (1-pvalue/\alpha)\}$; categoria (U para não diferencial e D para diferencial) e resultado da anotação automática.

O *score* serve como um parâmetro de ordenação dos *contigs* quanto ao nível de transcrição diferencial, ou seja, quanto maior é o *score*, mais diferencialmente expresso é o *contig*. A categoria (U ou D) foi decidida de acordo com o valor de p e o valor de α . Para *p-value* menor do que α , o *contig* é classificado como diferencialmente expresso. A Figura 8 mostra um exemplo de algumas linhas da saída do programa Kemp.

Para uma melhor visualização do nível de transcrição diferencial procedeu-se a normalização do número de contagens totais das bibliotecas provenientes dos diferentes estágios evolutivos. Os dados foram normalizados com base no número total de leituras da biblioteca de maior representatividade.

ID do cDNA	Ou	Op	Os	PE	p-value	alpha	score	category	Resultado da anotação automática
Emax_0001	0	0	0	2	0,02	0,03	3,13	D	Similar to clathrin assembly protein AP19
Emax_0002	0	12	1	13	0,00	0,03	10,00	D	Similar to GAD65/67 box helicase
Emax_0003	0	13	0	13	0,00	0,03	10,00	D	hypothetical protein
Emax_0004	0	13	0	13	0,00	0,03	10,00	D	hypothetical protein
Emax_0005	0	2	0	2	0,08	0,02	0,00	U	Similar to L-alanine dehydrogenase
Emax_0006	0	2	0	2	0,08	0,02	0,00	U	Similar to splicing factor, arginine/serine-rich 10 (transformer 2 homolog, Drosophila) variant
Emax_0007	0	2	0	2	0,08	0,02	0,00	U	hypothetical protein
Emax_0008	3	1	0	4	0,01	0,03	6,28	D	Similar to thymoproteptidase, putative
Emax_0009	1	0	1	2	0,75	0,02	0,00	U	Similar to alpha chain A, Crystal Structure of Toxoplasma Gondii Peptidyl-Prolyl Cis- Trans Isomerase
Emax_0010	2	0	0	2	0,04	0,02	0,00	U	hypothetical protein
Emax_0011	1	17	2	20	0,00	0,02	10,00	D	Similar to hypothetical protein PVX_115060
Emax_0012	4	0	0	4	0,00	0,03	9,42	D	Similar to vacuolar protein sorting 29, putative
Emax_0013	0	0	26	26	0,00	0,02	10,00	D	hypothetical protein
Emax_0014	0	0	4	4	0,19	0,03	0,00	U	Similar to hypothetical protein COPEUT_01588
Emax_0015	0	0	3	3	0,42	0,03	0,00	U	hypothetical protein
Emax_0016	0	0	2	2	0,52	0,02	0,00	U	hypothetical protein
Emax_0017	1	0	38	39	0,00	0,01	10,00	D	Similar to heat shock protein 90
Emax_0018	0	1	2	3	0,85	0,03	0,00	U	hypothetical protein
Emax_0019	0	0	2	2	0,52	0,02	0,00	U	hypothetical protein
Emax_0020	0	0	5	5	0,09	0,03	0,00	U	hypothetical protein
Emax_0021	0	0	6	6	0,08	0,03	0,00	U	hypothetical protein
Emax_0022	0	0	7	7	0,04	0,03	0,00	U	hypothetical protein
Emax_0023	0	0	5	5	0,09	0,03	0,00	U	hypothetical protein
Emax_0024	0	0	3	3	0,42	0,03	0,00	U	hypothetical protein
Emax_0025	0	0	2	2	0,52	0,02	0,00	U	hypothetical protein
Emax_0026	2	0	0	2	0,04	0,02	0,00	U	hypothetical protein
Emax_0027	2	0	0	2	0,04	0,02	0,00	U	hypothetical protein
Emax_0028	5	0	0	5	0,00	0,03	9,89	D	hypothetical protein
Emax_0029	50	47	1	98	0,00	0,01	10,00	D	hypothetical protein
Emax_0030	5	3	0	8	0,00	0,03	9,83	D	hypothetical protein
Emax_0031	1	0	1	2	0,75	0,02	0,00	U	hypothetical protein
Emax_0032	1	0	3	4	0,41	0,03	0,00	U	hypothetical protein
Emax_0033	0	0	4	4	0,19	0,03	0,00	U	hypothetical protein
Emax_0034	0	0	6	6	0,08	0,03	0,00	U	hypothetical protein
Emax_0035	0	0	2	2	0,52	0,02	0,00	U	hypothetical protein
Emax_0036	0	1	1	2	1,00	0,02	0,00	U	hypothetical protein
Emax_0039	4	0	0	4	0,00	0,03	9,42	D	Similar to hypothetical protein, conserved in Apicomplexan species
Emax_0040	0	1	0	1	0,41	0,03	0,00	U	hypothetical protein
Emax_0041	0	6	0	6	0,00	0,03	9,95	D	hypothetical protein
Emax_0042	0	4	0	4	0,00	0,03	8,74	D	Similar to PMCA-like calcium ATPase A1
Emax_0043	0	3	0	3	0,02	0,02	3,13	D	Similar to similar to TATA box binding protein-related factor 2 CG18009-PD, isoform 0 isoform 1
Emax_0044	0	8	0	8	0,00	0,03	10,00	D	hypothetical protein
Emax_0045	0	3	0	3	0,02	0,03	3,13	D	Similar to similar to cullin
Emax_0046	0	2	0	2	0,08	0,02	0,00	U	hypothetical protein
Emax_0047	0	0	4	4	0,19	0,03	0,00	U	hypothetical protein
Emax_0048	0	0	4	4	0,19	0,03	0,00	U	hypothetical protein
Emax_0049	0	0	2	2	0,52	0,02	0,00	U	Similar to hypothetical protein Chro.40217
Emax_0050	0	0	2	2	0,52	0,02	0,00	U	hypothetical protein
Emax_0051	0	0	9	9	0,01	0,03	5,81	D	Similar to hypothetical protein, conserved
Emax_0052	0	0	48	48	0,00	0,01	10,00	D	hypothetical protein
Emax_0053	0	1	2	3	1,00	0,02	0,00	U	hypothetical protein
Emax_0054	0	0	20	20	0,00	0,02	9,98	D	Similar to RNA-directed RNA polymerase II/III subunit
Emax_0055	0	0	5	5	0,09	0,03	0,00	U	hypothetical protein
Emax_0056	0	0	3	3	0,09	0,03	0,00	U	Similar to hypothetical protein PVX_083230
Emax_0057	0	0	6	6	0,08	0,03	0,00	U	hypothetical protein
Emax_0058	0	0	3	3	0,42	0,03	0,00	U	Similar to hypothetical protein BBOV_I1003090
Emax_0059	0	0	8	8	0,03	0,03	2,12	D	hypothetical protein
Emax_0060	0	0	35	35	0,00	0,01	10,00	D	hypothetical protein
Emax_0061	0	0	4	4	0,19	0,03	0,00	U	hypothetical protein
Emax_0063	1	0	1	2	0,75	0,02	0,00	U	hypothetical protein
Emax_0066	0	0	1	1	1,00	0,01	0,00	U	hypothetical protein
Emax_0067	0	0	6	6	0,08	0,03	0,00	U	hypothetical protein
Emax_0068	0	0	2	2	0,52	0,02	0,00	U	Similar to calcium-dependent protein kinase
Emax_0069	0	18	237	255	0,00	0,00	10,00	D	Similar to hypothetical protein, expressed
Emax_0070	0	0	6	6	0,08	0,03	0,00	U	hypothetical protein
Emax_0071	0	0	2	2	0,52	0,02	0,00	U	hypothetical protein
Emax_0072	0	0	6	6	0,08	0,03	0,00	U	hypothetical protein
Emax_0073	0	0	2	2	0,52	0,02	0,00	U	hypothetical protein
Emax_0074	0	0	2	2	0,52	0,02	0,00	U	Similar to AF288603_1 cAMP-dependent protein kinase regulatory subunit
Emax_0075	0	0	2	2	0,52	0,02	0,00	U	hypothetical protein
Emax_0076	0	0	3	3	0,42	0,03	0,00	U	hypothetical protein
Emax_0077	0	0	3	3	0,42	0,03	0,00	U	Similar to GNAT family histone acetyltransferase GCN5-B

Figura 8- Saída em modo texto gerado pelo programa Kemp. A primeira linha representa um cabeçalho contendo informações relativas a cada *contig* como: identificador do cDNA, número de leituras obtidas para cada estágio evolutivo, oocistos não esporulados (Ou), oocistos parcialmente esporulados (Op) e oocistos esporulados (Os); PE (poder de expressão); valores de p, α , score, categoria (U para não diferencialmente expresso e D para diferencialmente expresso) e resultado da anotação automática (programa BLAST).

3.17 Heat Map

Para uma melhor visualização do perfil de expressão gênica, os resultados obtidos foram inseridos no programa MeV (MultiExperiment Viewer), versão 4.6 que dentre várias análises, agrupa sequências utilizando diferentes correlações de distância e mostra o perfil de expressão na forma gráfica (*heat map*).

Para tanto, foi utilizado o algoritmo de agrupamento hierárquico (HCL) que agrupa as sequências em função da similaridade do perfil de expressão dos genes e/ou das amostras. No presente trabalho, a análise foi realizada empregando a correlação de Pearson.

3.18 Seleção e análise dos 100 reconstruídos diferencialmente expressos mais abundantes

Para análise dos genes diferencialmente expressos foram escolhidos 100 reconstruídos diferencialmente expressos que apresentavam o maior poder de expressão (PE) (item 3.16), ou seja, que eram os mais abundantes.

3.18.1 Busca de similaridade de transcritos de *E. maxima* com sequências de outros parasitas do Filo Apicomplexa

Com a finalidade de conhecer e identificar possíveis genes e proteínas conservadas entre as espécies do Filo Apicomplexa, os transcritos selecionados (diferencialmente expressos mais abundantes) foram submetidos à análise de busca de similaridade utilizando o programa BLAST. Como banco de dados foram empregadas as bases: PlasmoDB (<http://plasmodb.org>), CryptoDB (<http://cryptodb.org>) e ToxoDB (<http://toxodb.org>). Somente foram considerados semelhantes os alinhamentos cujos *e-values* eram inferiores a $1e-06$.

Os transcritos selecionados também foram comparados às sequências ORESTES de *E. acervulina* geradas pelo laboratório, dentro do projeto ORESTES (www.coccidia.icb.usp.br) e de *E. tenella* geradas pelo grupo e por terceiros (Sanger instiyute, UK; Universiti Kebangsaan Malaysia; Washington University, USA; USDA, USA), as quais foram montadas em conjunto com as nossas leituras (Projeto ORESTES de *E. tenella*). Para isto, foram realizadas buscas de similaridade por TBLASTX observando os alinhamentos e os resultados da anotação automática. Somente foram considerados semelhantes os alinhamentos cujos *e-values* eram inferiores a $1e-06$. Para os transcritos que eram similares entre si foi realizado um estudo comparativo dos perfis de expressão gênica das diferentes fases da esporulação. Esta última análise, além de ser empregada para os 100 transcritos diferencialmente expressos mais abundantes, foi também realizada para os genes diferencialmente expressos estudados por qPCR e que não estavam dentro deste grupo.

3.18.2 Análise no programa Threader 3 (Protein Fold Recognition by Threading)

Com a finalidade de inferir possíveis funções das proteínas hipotéticas (similaridade acima 1e-06 ou resultado negativo no programa BLAST) presentes nos 100 reconstruídos diferencialmente expressos mais abundantes, as sequências nucleotídicas foram traduzidas e submetidas à análise pelo programa Threader 3 (Jones, 1998).

O THREADER 3 é um programa capaz de prever o padrão de enovelamento de uma dada proteína comparando-a à estrutura de proteínas conhecidas. Esta premissa se baseia no fato de que duas proteínas de estrutura semelhante podem apresentar função similar. Depois de predita, a estrutura tridimensional da proteína alvo é comparada à de proteínas conhecidas depositadas no banco de dados RCSB PDB (*The Protein Data Bank*, <http://www.pdb.org/pdb/home>).

A análise foi empregada utilizando os parâmetros *default* do programa *Threader 3* (<http://bioinfadmin.cs.ucl.ac.uk/downloads/threader/manual.pdf>). O arquivo produzido pelo programa foi gerado em formato CSV, exportado para o programa Excel e em seguida analisado.

Para a interpretação dos dados, as proteínas foram avaliadas individualmente analisando-se os valores do *Z-score* (obtido após a análise da energia de pareamento entre a proteína alvo e o modelo alinhavado, energia da proteína nativa e energia de solvatação) e os valores individuais da energia nativa da proteína alvo e da energia do modelo alinhavado (*threaded*). Só foram consideradas as proteínas que apresentavam *Z-score* acima de 2,7 (índice mínimo de confiabilidade sugerido pelo manual do programa (<http://bioinfadmin.cs.ucl.ac.uk/downloads/threader/>) e que tinham um valor de energia da proteína nativa maior que o valor de energia do pareamento da proteína alvo com o modelo alinhavado.

3.19 PCR em tempo real

A técnica de PCR em tempo real (qPCR – PCR quantitativo) foi empregada com o objetivo de validar os resultados obtidos por ORESTES e foi empregada para um pequeno conjunto de genes. Considerando que ORESTES não é uma metodologia quantitativa, para validar os resultados, comparamos apenas os perfis de transcrição gênica observados nas duas técnicas, para as três fases da esporulação.

Foram selecionados como genes controle (Tabela 2) os transcritos reconstruídos de ORESTES que apresentavam *p-value* igual/ou próximo de 1 (hipótese nula verdadeira) e perfil de expressão semelhante entre os três estágios de esporulação (oocistos não esporulados, parcialmente esporulados e esporulados).

A seleção de genes diferencialmente expressos (Tabela 3) foi realizada a partir de transcritos reconstruídos presentes em um ou mais estágios conforme os seguintes critérios, estar preferencialmente entre os 100 *contigs* diferencialmente expressos mais abundantes, valor de *p* igual a zero, ter similaridade à proteínas conhecidas de parasitas do gênero *Eimeria* spp. e/ou do Filo Apicomplexa ou à proteínas hipotéticas conservadas entre o filo Apicomplexa.

A partir das sequências reconstruídas, pares de *primers* foram desenhados com auxílio do programa *OLIGO 4.0 Primer Analysis Software* (autoria de Wojciech Rychlik). Para tanto, diferentes características dos oligonucleotídeos, tais como temperatura de fusão (T_m , Temperatura de *melting*), complementaridade entre o par de *primers*, falsos sítios de pareamento, formação de estruturas tipo grampo de cabelo (*hairpins*), entre outros, foram analisados. O tamanho das sequências a serem amplificadas foi limitado a 400 pares de bases (Tabela 2 e 3).

Tabela 2- Primers testados em experimentos de PCR em tempo real para a escolha do gene controle (*housekeeping gene* constitutivamente expresso).

Nome	Sequência	Produto	Tamanho do Amplicon (pb)
Emax_0088_F	GGG GCT TCG TCG TTT GAT	Proteína hipotética	189
Emax_0088_R	AGG AGT GCG GTT TGG GAG		
Ema_0480_F	ACC TCT CTT GTT GTG TTG TTG C	Proteína hipotética	282
Emax_0480_R	GTT GTT GTT GTT GCT GCT GC		
Emax_0608_F	CCG AAA GCC AAA AGC CAA	Proteína similar a GH15714 (proteína relacionada à ubiquitina)	152
Emax_0608_R	GCT GGA GGA CGG TAG GAC AC		
Emax_0774_F	AAC ACC CGC AGC ATC CAC	Proteína hipotética	228
Emax_0774_R	GCC CCT ACT TGC CTG CCT		
Emax_0823_F	TGC TTG TCT TCT CCT GTT GAT G	Proteína similar à proteína 2g4 (fator associado à proliferação)	145
Emax_0823_R	TGC GAA GTC ACC AAA ATC ATA G		
Emax_0905_F	CAG CGG CAG CAA CGA AAA	Proteína hipotética	317
Emax_0905_R	GGC AGG GAG GTT TGA AGG AA		
Emax_0953_F	GCC AAA TCA GAA GCC GCA AG	Proteína hipotética	161
Emax_0953_R	CTC CGC CCT ACC GTT ATC CA		
Emax_1079_F	AAG CAA AGA GAC GCC ACG	Proteína hipotética	100
Emax_1079_R	AAG CAG CAG CAG CAG GAA		

Tabela 3- *Primers* utilizados para a validação da expressão gênica diferencial em experimentos de PCR em tempo real.

(Continua)

Nome	Sequência	Produto	Tamanho do Amplicon (pb)
Emax_0017_F	CTT GCT GCT GTG GAA CCG	Proteína de choque térmico 90	256
Emax_0017_R	TCG TGA AGG GTG TCG TCG		
Emax_0130_F	ATG GAG GTT GCT GAC TGT GC	Antígeno não transmembranar	240
Emax_0130_R	TTC AGC AGA GGG AGG AGA CA		
Emax_0208_F	CGT CAC CTT TCC CCC ATC	Precursor de proteína de micronema	209
Emax_0208_R	CCC TGC CTT CAA CTT CGG		
Emax_0247_F	TGA CAG AGA TTT TCG GCA GG	Proteína precursora de SERPINA	210
Emax_0247_R	TGG ATT GGT GTC TTT GGC AT		
Emax_0264_F	TCT GTT GTC GTG GGG TTC TC	Proteína putativa não caracterizada	367
Emax_0264_R	GAA AGC AGG CAC CAT CAA AG		
Emax_0357_F	TCG CTC ATC TTC CAG GCA	Dinamina	137
Emax_0357_R	GAC AAT GGA ACG GAC GCA		
Emax_0639_F	TTC GGG TCT GCT GCT CCA	Aspartil proteinase (eimepsina)	272
Emax_0639_R	TAT CGT GGA AAT CGG CGG		
Emax_0859_F	GCC CGC CAT ACC CAA CTT	Proteína hipotética	223
Emax_0859_R	TGC CAA GCC GAT TCC TGA		
Emax_0955_F	TCT CCT GCT GCG TCG TCG	Aminopectidase	140
Emax_0955_R	GGC GGT GGA AAG GGC ATA		

Tabela 3- *Primers* utilizados para a validação da expressão gênica diferencial em experimentos de PCR em tempo real.

(Conclusão)

Nome	Sequência	Produto	Tamanho do Amplicon (pb)
E _{max} _0999_F	TTC GGT CTC GTT GGC TCC	Proteína rombóide	285
E _{max} _0999_R	CAA GGG TGA GGG CGA TGA		
E _{max} _1094_F	ACC AAG AAG CCC CAC GAC	Transidrogenase	173
E _{max} _1094_R	CCT CAA CCC ACC CCT GTC		
E _{max} _1166_F	TCC AAC CAA AGC GGC AAT	Proteína da família BT1	108
E _{max} _1166_R	GCC CGA TTA CGA CGA TGC		
E _{max} _1232_F	GCC AAA CGC AAC CCA GAG	Proteína de desenvolvimento de esporozoíto	208
E _{max} _1232_R	CTA TTT CCC CCC GTT CCC		

Para proceder ao q-PCR, a síntese de cDNA foi realizada com 10 ng de RNAm empregando Oligo(dT)₁₂₋₁₈ (Invitrogen), enzima RevertAid™ H Minus M-MuLV (Fermentas International Inc, Canadá), conforme instruções dos fabricantes.

Os experimentos de validação do gene controle e das amostras diferencialmente expressas foram realizados empregando-se amostras de RNA mensageiro de três extrações independentes, purificadas a partir de oocistos não esporulados (Ou), oocistos parcialmente esporulados (Op) e oocistos esporulados (Os).

As reações de qPCR foram realizadas em triplicata em um volume final de 20 µL empregando o kit ABSolute™ QPCR SYBR Green Mix (Thermo Fisher Scientific Inc., Waltham, Massachusetts, EUA), 10% da reação de cDNA e os diferentes pares de primers. Para a detecção da amplificação do produto utilizou-se o termociclador *Mastercycle* Eppendorf (modelo Realplex2/2, Eppendorf), empregando a seguinte ciclagem: 15 minutos a 95 °C para ativação da *Taq* DNA polimerase, seguido de 35 ciclos de 15 segundos a 95 °C, 30 segundos a 51 °C e 30 segundos a 72 °C. Em seguida, para a determinação da especificidade dos produtos de PCR e verificação da formação de dímeros de *primers*, foram feitas curvas de dissociação (curvas de *melting*) dos produtos amplificados. Esta curva consiste em etapas de 95 °C por 30 segundos, 60 °C durante 30 segundos e um aquecimento gradual de 60 °C até 95°C, durante 20 minutos. Para cada gene testado reações sem a adição de cDNA (controle negativo) foram também realizadas.

Para a padronização da concentração de cada *primer*, diferentes concentrações de *primers* (200nM, 300nM, 400nM e 500nM) foram testadas empregando um *pool* de cDNA de oocistos não esporulados, parcialmente esporulados e esporulados. Em seguida, a curva de concentração dos primers foi traçada determinando-se a menor concentração de cada *primer* na qual a eficiência da reação se mantinha.

Para todos os genes, a eficiência das reações de PCR foi também determinada, sendo aceita uma variação do índice de eficiência entre 1,60 e 2,1. Procedeu-se o cálculo da eficiência da reação de acordo com a fórmula $E=10^{(-1/a)}$ (Pfaffl, 2001). As amostras (*pool* de cDNA extraídos de oocistos não esporulados, parcialmente esporulados e esporulado) foram diluídas (amostra pura, 1/5, 1/25 e 1/625), submetidas à reação de PCR em tempo real calculando-se em seguida o valor de inclinação da reta e a eficiência da reação.

A análise dos dados foi realizada utilizando o software Realplex (Eppendorf). Para a determinação do Ct, os dados de fluorescência foram coletados na fase exponencial da reação. O *threshold* foi automaticamente definido pelo software Realplex seguindo os parâmetros *default* do programa.

Para a quantificação da expressão, optamos pelo método de análise relativa, onde a expressão dos genes selecionados foi normalizada em relação a um gene controle. Para tanto, os valores de Ct foram exportados para o programa Q-Gene (Muller et al., 2002) e o $2^{-\Delta Ct}$ foi calculado empregando a equação 2 que utiliza concomitantemente a média dos valores de Ct das triplicatas tanto para o gene referência quanto para o gene controle. Os dados foram exportados para uma planilha Excel e os valores de $2^{-\Delta\Delta Ct}$ foram calculados. Em seguida, os perfis de expressão gênica obtidos por ORESTES e PCR em tempo real foram comparados.

4 RESULTADOS

4.1 Propagação dos parasitas

Após a infecção e purificação dos oocistos de *E. maxima*, cepa H, foi obtido um rendimento médio de 10^7 oocistos/ave. A taxa média de esporulação dos oocistos foi de 75%.

Com a finalidade de obter diferentes fases da esporulação de oocistos de *E. maxima* foi realizada uma cinética de esporulação empregando diferentes tempos de incubação. Foi constatado que entre 22 e 24 horas do início da esporulação cerca de 70% dos oocistos estavam na fase de esporoblastos (Figura 9B), sendo este o período escolhido para obtenção dos oocistos parcialmente esporulados.

A Figura 9 exemplifica o que é um oocisto não esporulado (A), parcialmente esporulado (B) e esporulado (C) de *E. maxima*.

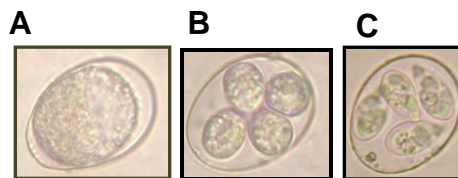


Figura 9- Fotomicrografias de oocistos de *E. maxima* em diferentes fases da esporulação: (A) oocisto não esporulado, (B) oocisto parcialmente esporulado e (C) oocisto esporulado.

4.2 Verificação de contaminação interespecífica

O DNA das amostras de oocistos utilizados para as infecções das aves, assim como o DNA dos oocistos recém-purificados foram submetidos a um ensaio de PCR multiplex capaz de detectar simultaneamente as sete espécies de *Eimeria* que infectam a galinha doméstica.

Em todos os ensaios de PCR multiplex realizados apenas uma única banda de 272 pb foi observada, este tamanho é compatível com o marcador SCAR específico de *E. maxima* (Fernandez et al., 2003a), como o exemplificado na Figura 10. O fato de não terem sido observadas bandas de outros tamanhos compatíveis

com as demais espécies de *Eimeria* é um indicativo de que não ocorreram contaminações interespecíficas das amostras de *Eimeria maxima* utilizadas.

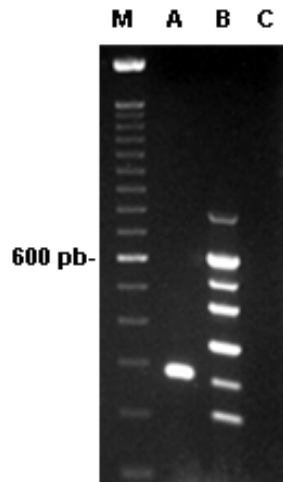


Figura 10- Resultado do ensaio de PCR *multiplex* para detecção das sete espécies de *Eimeria* de galinha doméstica. A reação foi realizada com: (A) DNA de oocistos de *E. maxima* obtidos após a purificação; (B) amostra mista de DNA contendo sete espécies de *Eimeria*; (C) controle negativo; (M) Marcador de tamanho molecular tipo escala de 100 pb (Invitrogen).

4.3 Avaliação da qualidade das amostras de RNA mensageiro

Amostras de RNA mensageiro de alta qualidade e livre de bandas ribossômicas, conforme exemplificado na Figura 11, foram utilizadas para a construção das bibliotecas de ORESTES e para os ensaios de qPCR.

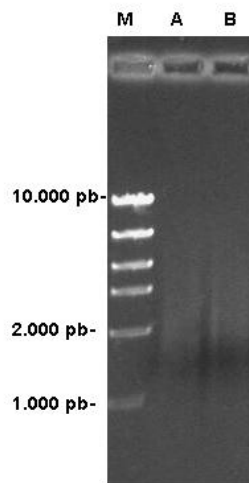


Figura 11- Eletroforese em gel de agarose do RNA mensageiro purificado e tratado com DNase RQ1. (A) RNA mensageiro extraído a partir de oocistos não esporulados, (B) RNA mensageiro extraído a partir de oocistos parcialmente esporulados de *Eimeria maxima* (M) Marcador de tamanho molecular *High Mass Ladder* (Invitrogen).

Além da visualização das amostras de RNA mensageiro em gel de agarose, a integridade do RNA mensageiro foi verificada pela amplificação de um transcrito de alto peso molecular de um gene codificante para a proteína TFP-250 (Genbank: AY239227) de *Eimeria maxima*. Após o ensaio de RT-PCR, para todos os RNAs utilizados foi observada uma banda de 900 pb compatível com o fragmento do gene de TFP-250, conforme exemplificado na Figura 12. Estes resultados indicam de forma indireta que as amostras de RNA mensageiro utilizadas apresentavam-se integras e, portanto, puderam ser utilizadas nas etapas seguintes deste trabalho.

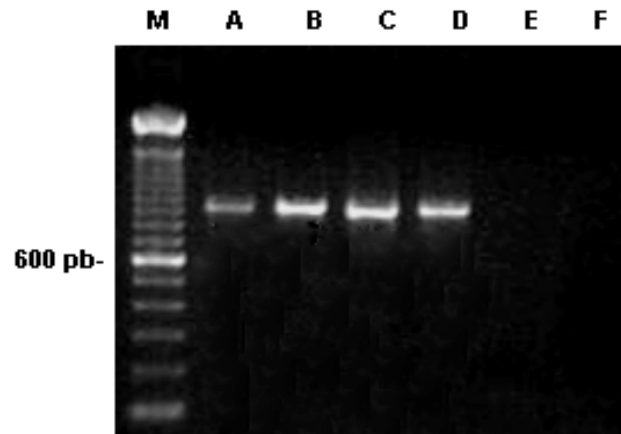


Figura 12- Resultado do ensaio de PCR de controle de integridade do RNA mensageiro de *E. maxima* (A) amostra de cDNA obtida a partir de RNA mensageiro de oocistos não esporulados, (B) de oocistos parcialmente esporulados, (C) de oocistos esporulados, (D) controle positivo da reação de PCR (E) controle negativo da reação de transcriptase reversa, (F) controle negativo da reação de PCR e (M) Marcador de tamanho molecular do tipo escala de 100 pb (Invitrogen).

Para verificar a presença de contaminação com DNA genômico, as amostras de RNA mensageiro purificado foram submetidas a uma reação de PCR capaz de detectar a existência de DNA genômico de *E. maxima*. Para todas as amostras, conforme exemplificado na Figura 13, não foi observada a presença de banda compatível (272pb) com o marcador SCAR espécie-específico de *E. maxima* que não é codificado (Fernandez et al., 2003b). Estes resultados indicam que as amostras de RNA mensageiro purificado estavam livres de contaminação por DNA genômico.

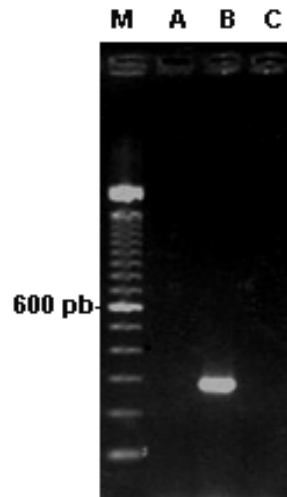


Figura 13- Resultado do ensaio de PCR para a verificação de contaminação de RNA mensageiro com DNA genômico de *E. maxima*. (A) RNA mensageiro de oocistos não esporulados (B) Controle positivo realizado empregando DNA de *E. maxima*, (E) Controle negativo da reação e (C) Marcador de tamanho molecular do tipo escala de 100 pb (Invitrogen).

4.4 Construção das bibliotecas de ORESTES

4.4.1 Obtenção e manipulação dos perfis de cDNA

Foram gerados 438 perfis de oocistos de *Eimeria maxima*, dos quais 75 foram descartados por terem pouca quantidade de material amplificado ou por apresentarem bandas bem definidas o que poderia gerar, neste último, bibliotecas com alta redundância.

Dos 363 perfis, 64 foram provenientes de oocistos não esporulados, 79 de oocistos parcialmente esporulados e 220 de oocistos esporulados.

A Figura 14 ilustra dois perfis de boa qualidade, cujo material amplificado apresentou-se como um rastro de fragmentos de DNA.

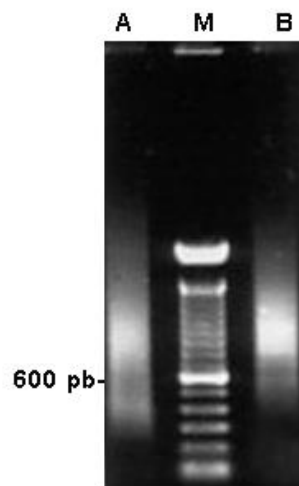


Figura 14- Eletroforese em gel de agarose 1,0% mostrando os perfis de cDNA das bibliotecas Mx093 (1) e Mx094 (2) provenientes de oocistos esporulados de *Eimeria maxima*. (M) Marcador de tamanho molecular de escala tipo 100 pb (Invitrogen).

4.4.2 Clonagem, transformação bacteriana e PCR de colônia

Perfis de tamanho entre 400 pb a 1000 pb foram selecionados, excisados do gel, purificados e ligados ao vetor pGEM-T-Easy (Promega). Após a clonagem, os insertos foram amplificados por PCR de colônia. A Figura 15 mostra algumas reações de PCR de colônia obtidas.

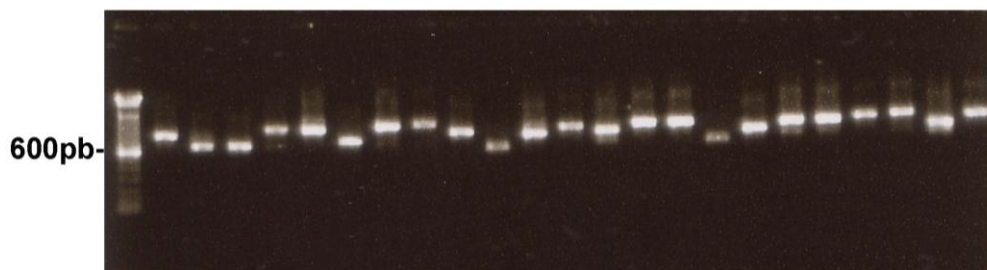


Figura 15- Produto da reação de PCR de colônia a partir de sequências clonadas da biblioteca de oocistos não esporulados, MxOu02601. (M) Marcador de tamanho molecular de escala tipo 100 pb (Invitrogen).

4.5 Sequenciamento de ESTs do tipo ORESTES de *Eimeria maxima*

Foram sequenciadas 20.149 leituras de DNA complementares aos RNAs mensageiros dos diferentes estágios da esporulação de *Eimeria maxima*, das quais 4.224 sequências eram provenientes de oocistos não esporulados (Ou), 3.840 de oocistos parcialmente esporulados (Op) e 12.085 de oocistos esporulados (Os).

4.6 Pré-processamento das sequências ORESTES de *Eimeria maxima*

Os eletroferogramas gerados foram submetidos a um pipeline de pré-processamento seriado automático, empregando-se o sistema EGene. A Tabela 4 mostra o número de sequências geradas, filtradas e aceitas.

Tabela 4- Pré-processamento das leituras ORESTES de *Eimeria maxima*.

Discriminação	Número de sequências	% de sequências
Sequências geradas	20.149	(100%)
Filtradas por qualidade	2.973	(14,76%)
Filtradas por tamanho	172	(0,85%)
Mitocôndria	5	(0,02%)
Apicoplasto	9	(0,04%)
Ribossomo	1.671	(8,29%)
Sequências repetitivas	42	(0,21%)
Bactéria	0	(0%)
Galinha	0	(0%)
Humano	23	(0,11%)
Sequências aceitas	15.254	(75,71%)

Analisando a Tabela 4 observa-se que a maior parte das sequências não aceitas (14,76%) foi filtrada por não apresentar boa qualidade. A contaminação das

sequências foi baixa, a maior ocorrência foi ao redor de 8% que era proveniente de sequências ribossômicas. Ainda na Tabela 4, observa-se que a taxa de aceitação das sequências foi de 75,71% (15.254 sequências), o que pode ser considerado um bom resultado.

Das sequências aceitas, 3.118 eram originárias de bibliotecas construídas a partir de oocistos não esporulados (Ou), 3.215 de oocistos parcialmente esporulados (Op) e 8.921 de oocistos esporulados (Os) (Tabela 5).

4.7 Agrupamento das sequências ORESTES de *E. maxima*

Após a montagem das sequências realizadas no programa CAP3 empregando os parâmetros (p 90 e y 250) foi obtido um total de 3.298 eventos sendo 1.207 (36,6%) *contigs* e 2.091 (63,4%) *singlets*.

Dos *contigs* obtidos, 324 apresentavam leituras provenientes de oocistos não esporulados, 417 de oocistos parcialmente esporulados e 860 de oocistos esporulados (Tabela 5). Dos *singlets*, 442 eram provenientes de oocistos não esporulados, 471 de oocistos parcialmente esporulados e 1178 de oocistos esporulados (Tabela 5).

Tabela 5- Número de eventos e redundância (em porcentagem) observados após a montagem das sequências ORESTES de oocistos não esporulados (Ou), oocistos parcialmente esporulados (Op) e oocistos esporulados (Os) de *Eimeria maxima*.

Bibliotecas	Estágio			Total
	Ou	Op	Os	
Sequências	3118	3215	8921	15254
Contigs	324	417	860	1207
Singlets	442	471	1178	2091
Redundância (%)	76,01	72,37	77,15	78,37

Ainda na tabela 5, observamos que o número total de *contigs* obtidos (1207) não é equivalente ao valor da soma dos *contigs* que apresentaram leituras de cada estágio estudado, pois um mesmo *contig* pode conter leituras provenientes de mais de um estágio. O mesmo já não acontece em relação aos *singlets*.

Também nota-se na Tabela 5 que a redundância média obtida nas bibliotecas foi de 75%, valor este bastante alto.

O número de leituras presentes em cada *contig* pode ser visualizado na Figura 16. Observa-se que a maior parte dos *contigs* (98,34%) possuía menos do que 100 leituras, sendo que a maioria (85,5%) tinha até 10 leituras.

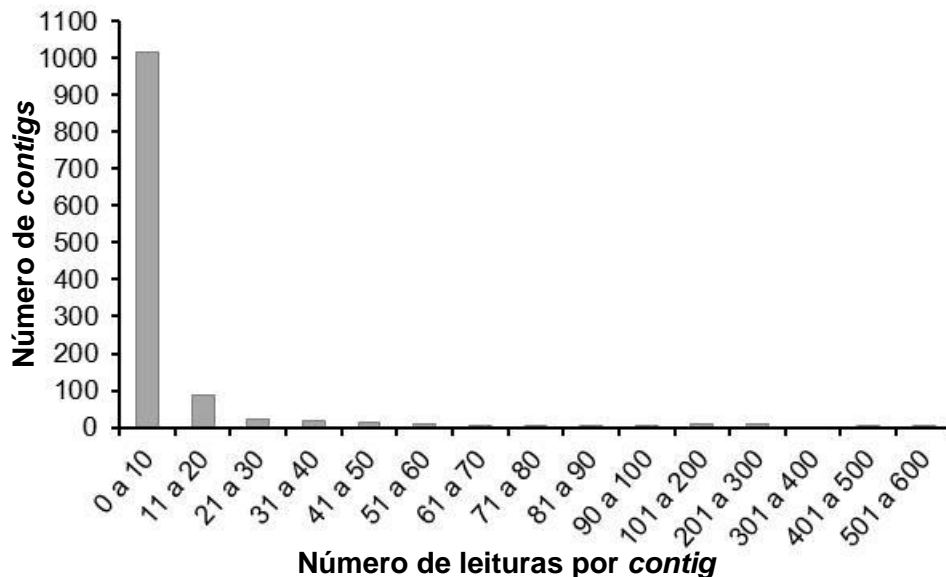


Figura 16- Histograma do número de leituras em cada *contig* em relação à sua frequência nas bibliotecas de *Eimeria maxima*.

O tamanho médio dos *contigs* foi de 558,32 nucleotídeos, a maior parte (44%) destes tinha tamanho variando entre 300 a 600 nucleotídeos (Figura 17). Já os *singlets*, tinham um tamanho médio de 318,49 nucleotídeos (dados não mostrados).

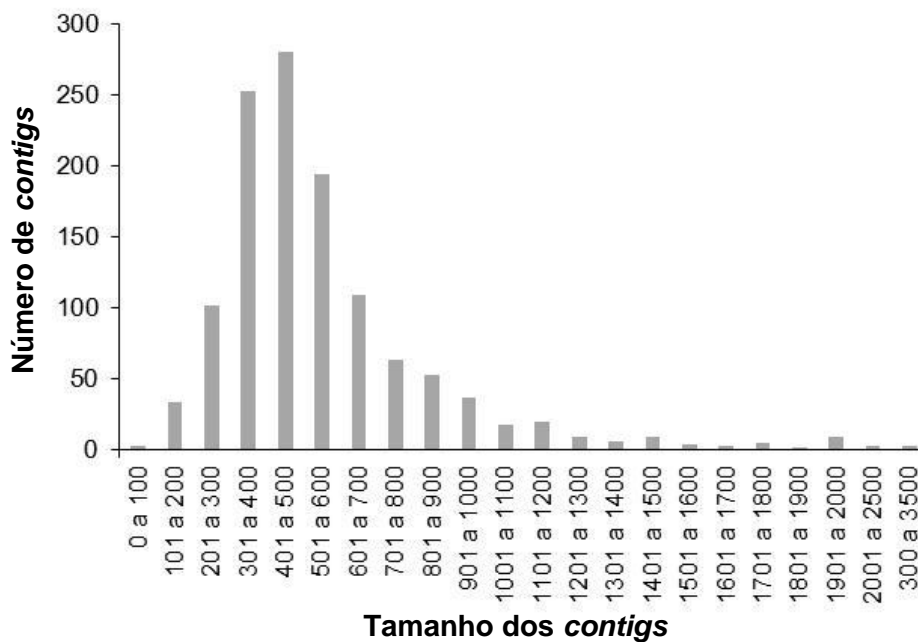


Figura 17- Histograma de distribuição do tamanho dos *contigs* (em nucleotídeos) em relação à sua frequência nas bibliotecas de *Eimeria maxima*.

4.8 Análise da posição relativa dos transcritos ORESTES em cDNAs completos de *E. maxima*

Um conjunto de cDNAs completos (*full length*) de *Eimeria maxima* (11 genes) presentes na base pública Genbank foram comparados com os *contigs* ORESTES gerados. Segue a lista dos referidos genes: antígenos de superfície de gametócitos, *gam82* (AY179510.2) e *gam56* (AY129951.2); proteína TFP-250 (AY239227); enzima lactato desidrogenase LDH (AY143390.1); duas fosfatidilinositol 4-fosfato 5-quinase, PIP5K (GQ293357.1 e GQ293358.1); proteína quinase cGMP (AF465543.1); proteína de choque térmico HSP70 (Z46964); proteína quinase com domínio de calmodulina (Z71756.1); antígeno de superfície *emp100* (M99058.1) e proteína de ligação à cadeia pesada de imunoglobulina (Z66491.1).

Após a realização de busca de similaridade dos cDNAs completos com os transcritos ORESTES, verificou-se que apenas 7 destes tinham alta similaridade (e *value* 0,0) com as sequências dos transcritos reconstruídos, Tabela 6.

Tabela 6- Lista dos cDNAs completos depositados no Genbank que eram similares (*e-value* 0,0) aos transcritos ORESTES de *Eimeria maxima*.

Genes completos (Genbank)	ID dos transcritos ORESTES
Antígeno de superfície (emp100)	Emax_1206
Proteína de choque térmico HSP70	Emax_1044
Fosfatidilinositol 4-fosfato 5- quinase (PIP5K)	Emax_0653
Fosfatidilinositol 4-fosfato 5- quinase (PIP5K)	Emax_1678
Proteína TFP-250	Emax_0767
Proteína quinase cGMP	Emax_0757
Proteína quinase com domínio de calmodulina	Emax_3168

Conforme podemos observar na Figura 18, a maioria dos transcritos mapearam na região central próxima extremidade 3' dos cDNA completos analisados. Após a busca de ORFs nos genes completos observou-se que a maioria (71,4%) dos transcritos mapeava na região codificante do gene correspondente.

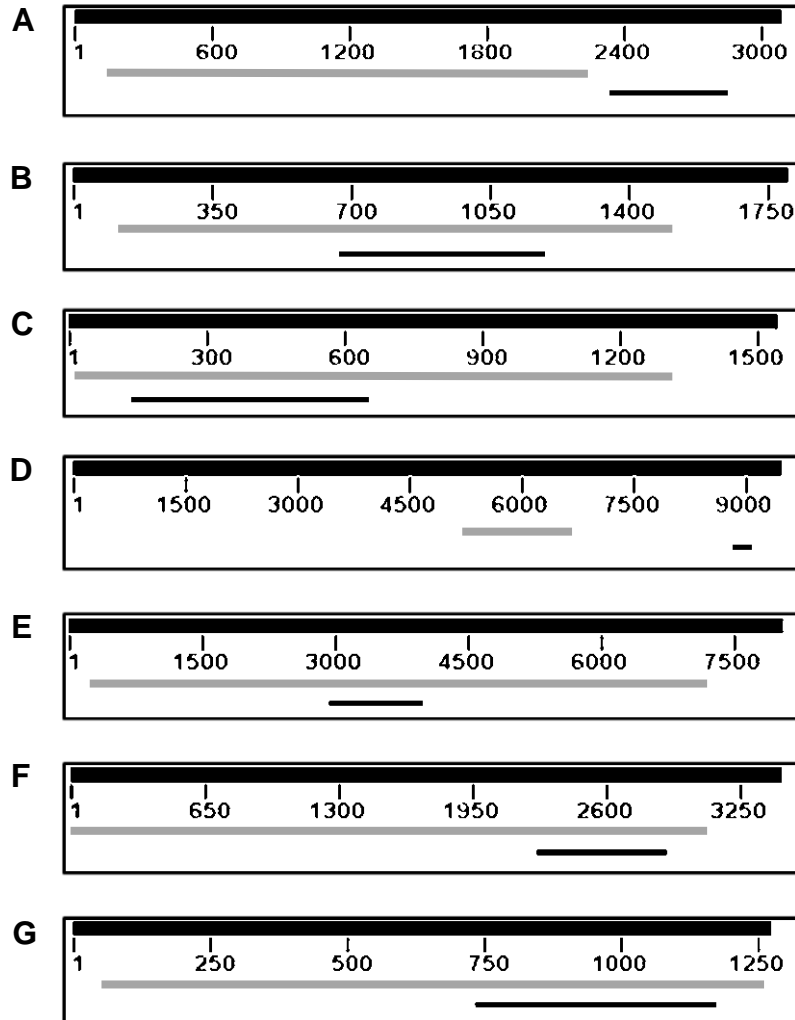


Figura 18: Alinhamento das seqüências nucleotídicas dos genes completos depositados no Genbank (barra de cor preta, mais larga) com as seqüências ORESTES similares (barra de cor preta, mais fina) e ORFs preditas (barra de cor cinza) A figura (A) representa o antígeno de superfície, (B) proteína de choque térmico HSP70, (C) fosfadilinositol 4-fosfato 5-quinase (PIP5K), (D) fosfadilinositol 4-fosfato 5-quinase (PIP5K), (E) proteína TFP-250, (F) proteína quinase cGMP e (G) proteína quinase com domínio de calmodulina.

4.9 Mapeamento dos *contigs* contra o genoma de *Eimeria maxima*

O conjunto dos 1.207 *contigs* foi submetido a um mapeamento contra o banco de *contigs* do genoma de *E. maxima*, conforme descrito no item 3.12. Para obtenção de um melhor resultado, diferentes percentuais de similaridade e de alinhamento entre os transcritos ORESTES e os *contigs* do genoma foram avaliados. Em relação ao alinhamento, os valores foram estabelecidos em função do percentual do tamanho do *contig* que mapeava no genoma.

Tabela 7- Mapeamento dos *contigs* ORESTES no genoma de *Eimeria maxima* empregando diferentes parâmetros de alinhamento e de similaridade.

Alinhamento (%) *	Similaridade (%)	Número de <i>contigs</i> mapeados	Porcentagem de <i>contigs</i> mapeados
30	90	1038	85,99%
30	95	1027	85,08%
40	90	945	78,29%
40	95	933	77,30%
50	90	828	68,59%
50	95	815	67,52%
75	90	576	47,72%
75	95	569	47,14%
90	90	478	39,60%
90	95	474	39,27%
100	90	369	30,57%
100	95	369	30,57%
100	100	173	14,33%

* Percentual mínimo do tamanho dos transcritos ORESTES que alinharam com os *contigs* do genoma,

Analisando a Tabela 7 observou-se que apenas 173 (14,33%) transcritos mapeados eram idênticos (100% de alinhamento e de similaridade) aos *contigs* do genoma. Com a diminuição da estringência dos parâmetros, principalmente do tamanho mínimo de alinhamento, foi observado um aumento do número de *contigs* mapeados. Quando foi empregado um tamanho mínimo de alinhamento de 40% e

similaridade de no mínimo 95% foi observado que 77,30% dos *contigs* mapeavam no genoma.

4.10 Anotação automática

A anotação automática dos transcritos de *Eimeria maxima* foi realizada empregando o sistema EGene (Ferro, 2008).

Dos 3.298 eventos obtidos (1207 *contigs* e 2091 *singlets*), 2.990 (90%) apresentaram pelo menos uma ORF, sendo que destes, 1194 eram *contigs* e 1796 eram *singlets*. Nos 308 transcritos restantes não foi detectada nenhuma ORF, destes, 13 eventos eram *contigs* e 295 *singlets*.

Dos 1194 *contigs* de *E. maxima* anotados, 865 *contigs* (72%) apresentaram produtos de baixa similaridade ($E\text{-value} < 10^{-6}$) contra a base de dados não redundante (*nr*), sendo consideradas como proteínas hipotéticas, 90 (8%) eram similares a proteínas hipotéticas de outros organismos e 239 (20%) eram semelhantes à proteínas de função conhecida (Figura 19). Dos 1796 *singlets* totais 67,3% eram semelhantes a proteínas hipotéticas e 18,55% eram similares à proteínas de função conhecida ou à proteínas hipotéticas conservadas (dados não mostrados).

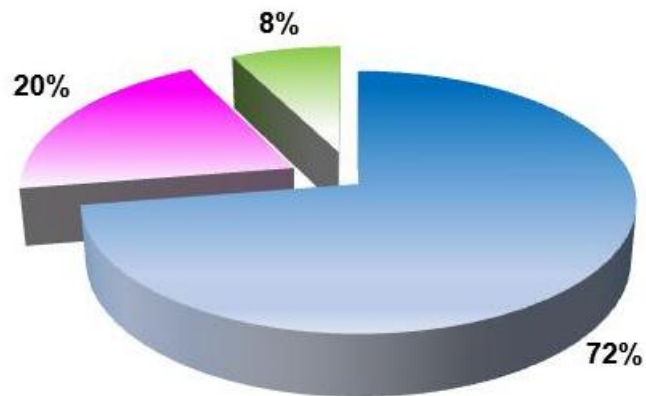


Figura 19- Distribuição dos *contigs* totais de *E maxima* quanto à anotação automática. proteínas hipotéticas (■), proteínas hipotéticas conservadas (■) e proteínas de função conhecida (■).

Dos 329 *contigs* que apresentaram similaridade com proteínas de função conhecida (239 *contigs*) ou proteínas hipotéticas conservadas (90 *contigs*), 74,2% (244) eram similares a proteínas descritas para o Filo Apicomplexa, dentre estes, 131 eram semelhantes à proteínas de parasitas da classe Coccidia e 46 à proteínas do Gênero *Eimeria* (Figura 20). Dentre os parasitas do Filo Apicomplexa, 68 transcritos eram similares a proteínas de *Plasmodium* spp, 25 de *Theileria* spp., 17 de *Babesia* spp., 48 de *Toxoplasma gondii*, 34 de *Cryptosporidium* spp., 1 de *Neospora caninum*, 1 de *Clyclospora cercopitheci* e 4 de *E. maxima*.

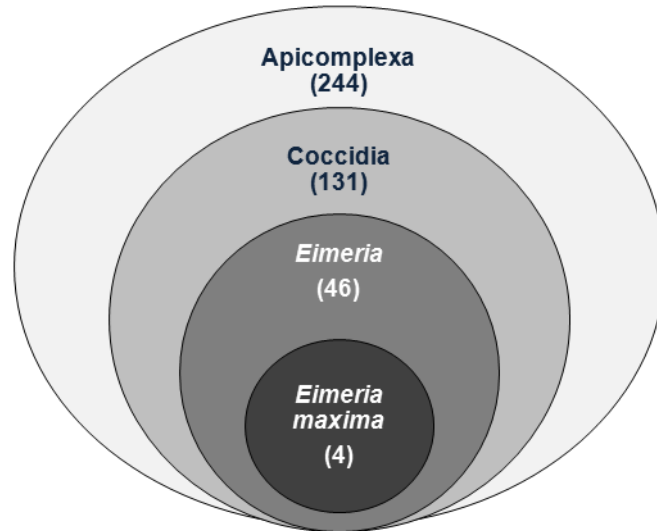


Figura 20- Diagrama de Venn mostrando a distribuição de produtos proteicos provenientes de ORESTES em relação ao Filo Apicomplexa, à classe Coccidia, ao Gênero *Eimeria* e à espécie *E.maxima*, após buscas de similaridade por BLAST em banco de dados não redundante.

Analisando os produtos proteicos semelhantes a proteínas do Gênero *Eimeria* (Figura 21), observamos que 12 proteínas são relacionadas ao complexo apical (10 proteínas de micronema, 1 de roptria e 1 proteína rombóide); 6 antígenos de superfície; 12 proteínas relacionadas ao metabolismo (5 proteínas quinases, 2 serino proteases, 1 proteína cálcio dependente, 1 aspartil proteinase, 1 proteína da família RB1, 1 manitol 1-fosfato desidrogenase e 1 transidrogenase); 10 outras proteínas com diferentes funções (3 proteínas de choque térmico, 2 proteínas de desenvolvimento de esporozoítos, 2 tubulinas, 1 histona, 1 proteína ribossômica e 1 fator depolimerizante de actina) e 6 proteínas hipotéticas conservadas.

Os produtos similares a proteínas de *Eimeria maxima* eram relacionadas a proteínas de micronema, sendo que, três destas eram similares a TFP-250 e uma à Etp100.

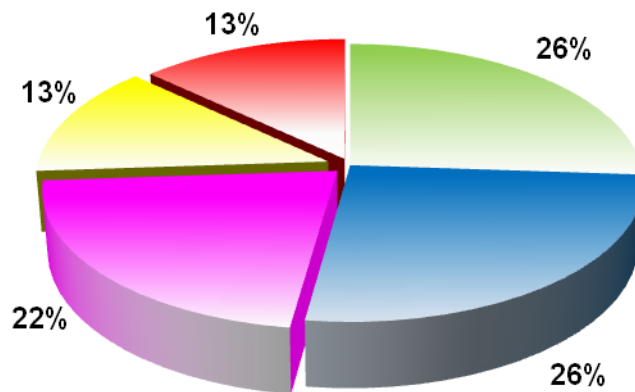


Figura 21- Distribuição dos produtos proteicos anotados automaticamente similares a proteínas do gênero *Eimeria*. Proteínas relacionadas ao metabolismo (■), proteínas do complexo apical (■), antígenos (■), demais proteínas (■) e proteínas hipotéticas conservadas (■).

4.11 Mapeamento dos termos de ontologia gênica (GO)

Dos 1194 eventos anotados, apenas 242 (20,3%) apresentaram resultados positivos contra as bases de dados do Interpro e puderam ser classificados quanto aos termos do GO (*Gene Ontology*). Os eventos que não foram categorizados receberam números de GO referentes às três ontologias gênicas principais. Os resultados de mapeamento GO para as três categorias estão mostrados a seguir: componente celular (Figura 22), processo biológico (Figura 23) e função molecular (Figura 24). As categorias com o maior número de representantes em componentes celulares estão relacionadas às células (6,2%), em processo biológico, ao metabolismo (10,1%), e em função molecular, à atividade catalítica (19,8%).

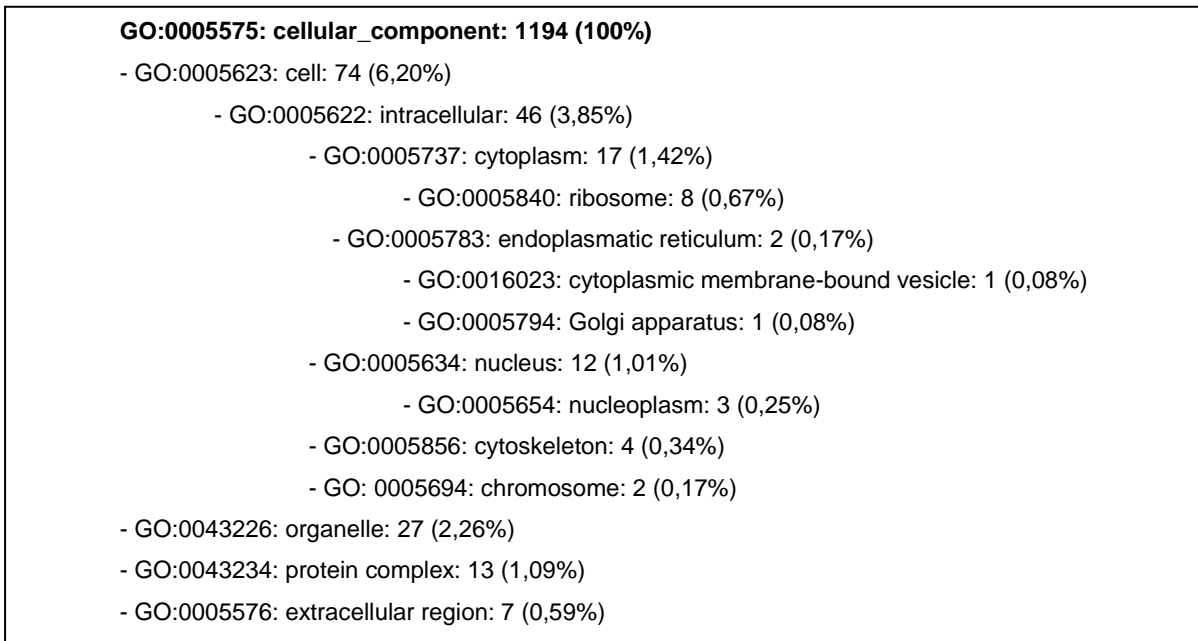


Figura 22- Gráfico acíclico de termos GO (ontologia gênica) das sequências proteicas codificadas pelos cDNAs reconstruídos a partir dos *contigs* de ORESTES de *E. maxima* para a categoria componentes celulares.

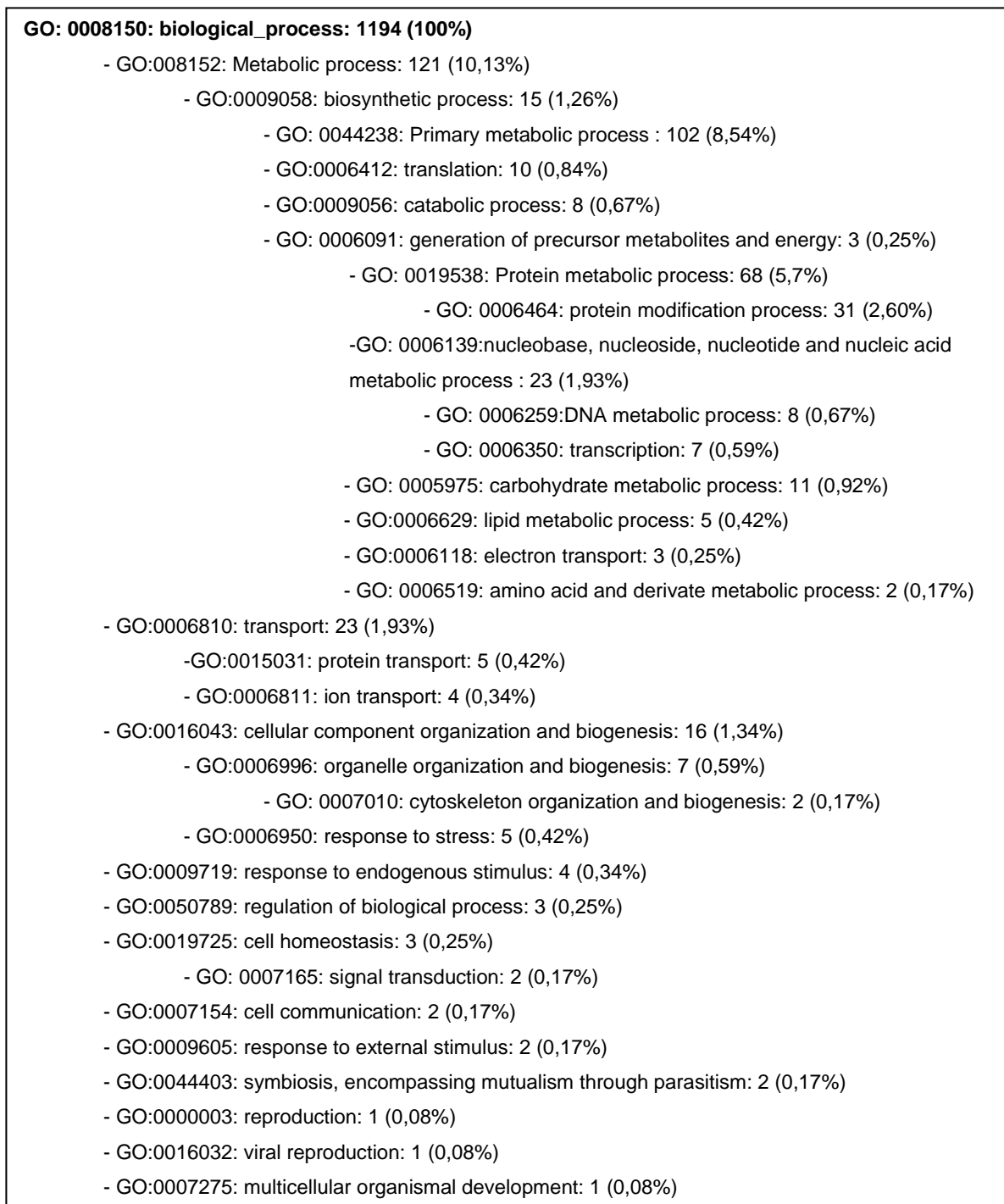


Figura 23- Gráfico acíclico de termos GO (ontologia gênica) das sequências proteicas codificadas pelos cDNAs reconstruídos a partir dos *contigs* obtidos de leituras de ORESTES de *E. maxima* para a categoria processos biológicos.

GO:0003674: molecular_function: 1194 (100%)

- GO:0003824: catalytic activity: 237 (19,85%)
 - GO:0016740: transferase activity: 39 (3,27%)
 - GO:0016301: kinase activity: protein binding: 24 (2,01%)
 - GO:0004672: protein kinase activity: 23 (1,93%)
 - GO: hydrolase activity: 36 (3,02%)
 - GO:0004518: nuclease activity: 4 (0,34%)
 - GO:0008233: peptidase activity: 14 (1,17%)
 - GO:0009055: electron carrier activity: 1 (0,08%)
- GO:0005488: binding: 152 (12,73%)
 - GO:0000166: nucleotide binding: 63 (5,28%)
 - GO:0003676: nucleic acid binding: 39 (3,27%)
 - GO:0003677: DNA binding: protein binding: 10 (0,84%)
 - GO:0003700: transcription factor activity: 3 (0,25%)
 - GO:0003723: RNA binding: 8 (0,67%)
 - GO:0008135: translation factor activity, nucleic acid binding: 2 (0,17%)
 - GO:0005515: protein binding: 24 (2,01%)
 - GO:0005102: receptor binding: 4 (0,34%)
 - GO:0008092: cytoskeletal protein binding: 2 (0,17%)
 - GO:0003779: actin binding: 2 (0,17%)
 - GO:0005509: calcium ion binding: protein binding: 11 (0,92%)
 - GO:0030246: carbohydrate binding: 2 (0,17%)
 - GO:0008289: lipid binding: 2 (0,17%)
- GO:0005198: structural molecule activity: 12 (1,01%)
- GO:0005215: transporter activity: 10 (0,84%)
- GO:0030234: enzyme regulator activity: 5 (0,42%)
- GO:0030528: transcriptional regulator activity: 3 (0,25%)
- GO:0004871: signal transducer activity: 2 (0,17%)
 - GO:0004872: receptor activity: 2 (0,17%)
- GO:0045182: translation regulator activity: 2 (0,17%)
- GO:0003774: motor activity: 1 (0,08%)

Figura 24- Gráfico acíclico de termos GO (ontologia gênica) das sequências proteicas codificadas pelos cDNAs reconstruídos a partir dos *contigs* obtidos de leituras de ORESTES de *E. maxima* para a categoria funções moleculares.

4.12 Northern Digital dos transcritos ORESTES de *E. maxima*

Com a finalidade de avaliar o perfil de expressão para os diferentes estágios estudados, os transcritos obtidos foram submetidos à análise digital da expressão empregando o *Northern Digital* (Audic e Claverie, 1997). Dos 1207 *contigs*, 73,5% (886) eram transcritos provenientes de somente de um estágio, 20,5% (248) de dois estágios e apenas 6% (73) dos três estágios estudados (Figura 25).

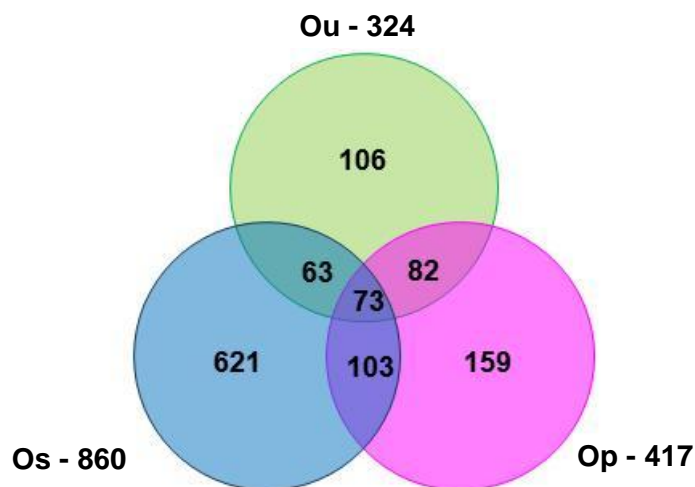


Figura 25- Diagrama de Venn mostrando a distribuição das sequências obtidas quanto à fase da esporulação: oocistos não esporulados (Ou), oocistos parcialmente esporulados (Op) e oocistos esporulados (Os).

4.13 Análise dos dados no Programa Simcluster

Com o objetivo de avaliar se os perfis de expressão gênica eram correlacionados entre si e se as relações observadas eram condizentes com a biologia do parasita, os transcritos obtidos foram agrupados hierarquicamente utilizando o programa Simcluster (Vencio et al., 2007). Diversos métodos de distância métrica, como Euclidiano, Aitchison, Pearson *Pseudo-Distance* e *Uncentered Pearson Pseudo-Distance*, foram empregados. Para tanto, foram empregados dados de expressão provenientes de oocistos obtidos no presente trabalho e dados de merozoítos de *E. maxima* gerados pelo grupo.

A análise dos resultados obtidos pelos diferentes métodos de distância mostrou que a distribuição dos agrupamentos em função dos estágios foram semelhantes e concordantes entre si (dados não mostrados). Abaixo segue uma figura exemplificando a análise pelo método de Aitchison (Figura 26).

Ao analisar o dendrograma, observamos que os oocistos estão num clado enquanto que os merozoítos estão presentes em outro clado mais externo. A fase de oocistos não esporulados e parcialmente esporulados está agrupada em um clado, mostrando uma maior relação entre estes.

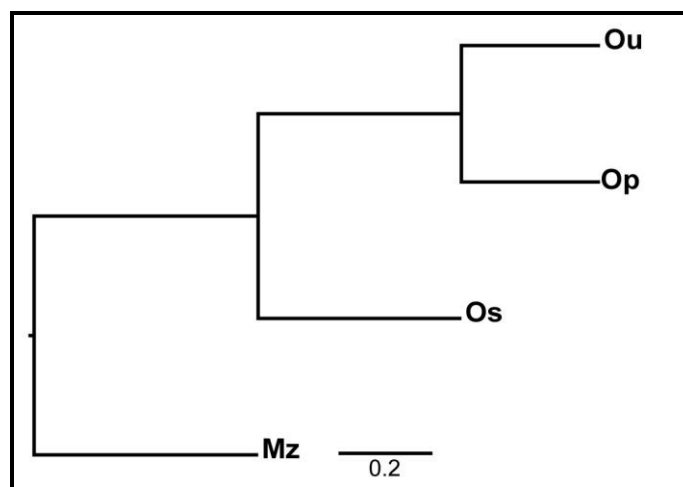


Figura 26- Dendrograma obtido pelo método de distancia de Aitchison a partir dos perfis de expressão gênica dos transcritos ORESTES de oocistos não esporulado (Ou) oocistos parcialmente esporulados (Op), oocistos esporulados (Os) e merozoítos (Mz) de *E. maxima*.

4.14 Análise estatística

Para identificar os *contigs* diferencialmente expressos, os transcritos reconstruídos foram submetidos à análise estatística empregando o programa Kemp (Varuzza, 2008). Este programa, além de gerar dados de *p-value*, possui um procedimento automático para o cálculo do nível crítico (α) para cada *contig* em particular. No presente trabalho, foi escolhido peso 4 para a hipótese nula ser verdadeira e peso 1 para hipótese nula ser falsa, desta forma, quando o *p-value* é menor do que α , o *contig* é automaticamente classificado como diferencialmente expresso.

Do conjunto dos 1207 *contigs*, 379 (31,4%) foram classificados como diferencialmente expressos e 828 (68,6%) como não diferencialmente expressos.

Quando analisamos o padrão de transcrição dos *contigs* diferencialmente expressos nas três fases da esporulação, observamos que há 15 diferentes perfis de expressão gênica (Tabela 8). Conforme podemos notar, a maioria dos transcritos está presente em um único estágio: 48 transcritos foram encontrados somente em oocistos não esporulados (Ou), 83 em oocistos parcialmente esporulados (Op) e 116 em oocistos esporulados (Os). Apenas 47 *contigs* (12,5%) foram observados nos três estágios da esporulação estudados.

Tabela 8- Distribuição dos transcritos ORESTES diferencialmente expressos em função dos perfis de expressão encontrados nas diferentes fases da esporulação: (Ou) oocistos não esporulados, (Op) oocistos parcialmente esporulados e (Os) oocistos esporulados.

Perfil de expressão	Número de transcritos encontrados
Ou	48
Ou>Op	25
Ou>Os	7
Ou>Op>Os	29
Ou>Os>Op	2
Op	83
Op>Ou	21
Op>Os	9
Op>Ou>Os	9
Op>Os>Ou	5
Os	116
Os>Ou	7
Os>Op	16
Os>Ou>Op	1
Os>Op>Ou	1

4.15 Análise comparativa da anotação automática de transcritos diferencialmente expressos com não diferencialmente expressos

Para os transcritos diferencialmente expressos, 376 *contigs* apresentaram pelo menos uma ORF, destes 260 eram proteínas hipotéticas, 77 proteínas de função conhecida e 39 proteínas hipotéticas conservadas (Figura 27, A). Dos *contigs* não diferencialmente expressos, 818 apresentaram pelo menos uma ORF, destes 605 eram proteínas hipotéticas, 151 proteínas de função conhecida e 62 proteínas hipotéticas conservadas (Figura 27, B). Conforme podemos observar, tanto para os genes diferencialmente expressos quanto para os não diferencialmente expressos, a maior parte das proteínas foi classificada como proteínas hipotéticas.

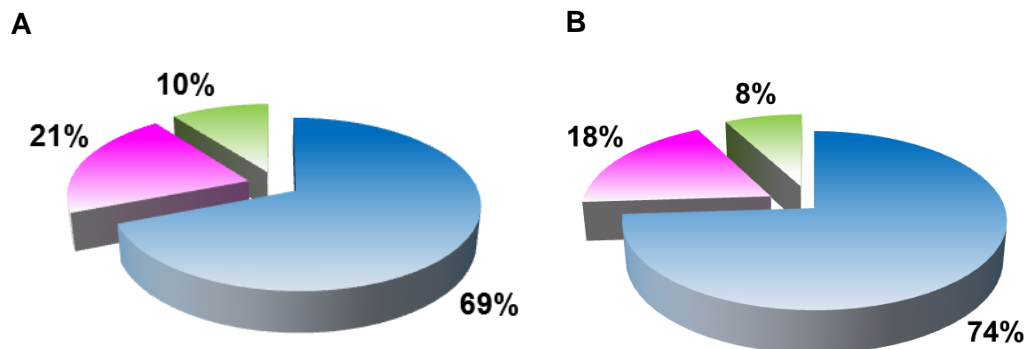


Figura 27- Distribuição dos *contigs* anotados automaticamente. Proteínas hipotéticas (■), proteínas hipotéticas conservadas (■) e proteínas de função conhecida (■). (A) *contigs* diferencialmente expressos e (B) *contigs* não diferencialmente expressos.

Nas Tabelas 9 e 10 consta respectivamente os resultados da anotação automática para os *contigs* diferencialmente e não diferencialmente expressos cujos produtos foram similares à proteínas já descritas.

Em relação aos transcritos diferencialmente expressos (Tabela 9), dentre as proteínas de função conhecida observamos que a maioria (34%) era relacionada ao metabolismo (aminopeptidase, proteínas quinases, cisteíno proteases, serpina, transportadores, entre outras), 5% relacionadas ao complexo apical (micronema, proteína de roptria, proteína rombóide); 4% a antígenos de superfície e o restante a proteínas de diferentes funções, tais como actina, proteína de desenvolvimento do esporozoítio, calmodulina, GTP binding, proteínas de choque térmico, entre outras.

Para os transcritos não diferencialmente, também, podemos observar que muitas proteínas (46%) estavam relacionadas ao metabolismo (proteínas quinases, desidrogenases, transportadores, entre outras); 3,3% relacionadas ao complexo apical (micronema, proteína de roptria); 1,3% a antígenos de superfície e as demais relacionadas a diferentes funções, tais como proteína de desenvolvimento do esporozoíto, calmodulina, tubulina, ubiquitina, proteínas de choque térmico, entre outras.

Tabela 9- *Contigs* diferencialmente expressos entre os estágios da esporulação de *E.maxima* que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), representado pela soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continua)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
Emax_1094	423	0,00E+00	1,40E-03	10,0	<i>Similar to transhydrogenase</i>	AAA29077.1	<i>Eimeria tenella</i>
Emax_0264	265	0,00E+00	2,20E-03	10,0	<i>Similar to unknown</i>	ABQ41429.1	<i>Eimeria tenella</i>
Emax_0526	262	0,00E+00	2,23E-03	10,0	<i>Similar to unknown</i>	ABQ41432.1	<i>Eimeria tenella</i>
Emax_0069	255	0,00E+00	2,29E-03	10,0	<i>Similar to hypothetical protein, expressed</i>	CAJ20353.1	<i>Toxoplasma gondii</i>
Emax_1190	220	0,00E+00	2,64E-03	10,0	<i>Similar to SDP_EIMBO Sporozoite developmental protein</i>	P42789	<i>Eimeria bovis</i>
Emax_0880	189	0,00E+00	3,06E-03	10,0	<i>Similar to unknown protein RB1-a</i>	AAC26122.1	<i>Eimeria acervulina</i>
Emax_0150	166	0,00E+00	3,47E-03	10,0	<i>Similar to zinc-finger protein, putative</i>	CAQ38960.1	<i>Plasmodium knowlesi</i>
Emax_0247	138	0,00E+00	4,15E-03	10,0	<i>Similar to SERPIN1 protein precursor</i>	CAI72624.1	<i>Eimeria tenella</i>
Emax_1166	135	0,00E+00	4,23E-03	10,0	<i>Similar to folate/biopterin transporter family protein</i>	EEA06295.1	<i>Cryptosporidium muris</i>
Emax_1232	110	0,00E+00	5,16E-03	10,0	<i>Similar to SDP_EIMBO Sporozoite developmental protein</i>	P42789	<i>Eimeria bovis</i>
Emax_0639	98	0,00E+00	5,78E-03	10,0	<i>Similar to aspartyl protease</i>	CAA80843.1	<i>Eimeria acervulina</i>
Emax_0142	92	0,00E+00	6,14E-03	10,0	<i>Similar to hypothetical protein, conserved in Apicomplexan species</i>	CAQ42141.1	<i>Plasmodium knowlesi</i>
Emax_0130	79	0,00E+00	7,12E-03	10,0	<i>Similar to non-transmembrane antigen</i>	AAO65977.1	<i>Toxoplasma gondii</i>
Emax_0152	79	0,00E+00	7,12E-03	10,0	<i>Similar to hypothetical protein PVX_090030</i>	XP_001615098.1	<i>Plasmodium vivax</i>
Emax_0435	76	0,00E+00	7,39E-03	10,0	<i>Similar to hypothetical protein MGL_0149</i>	XP_001732374.1	<i>Malassezia globosa</i>
Emax_0297	66	0,00E+00	8,47E-03	10,0	<i>Similar to hypothetical protein</i>	XP_980864.1	<i>Mus musculus</i>
Emax_0517	60	0,00E+00	9,29E-03	10,0	<i>Similar to cation-transporting ATPase</i>	ACB97673.1	<i>Eimeria acervulina</i>
Emax_1165	52	0,00E+00	1,07E-02	10,0	<i>Similar to hypothetical protein, conserved in Apicomplexan species</i>	CAQ38339.1	<i>Plasmodium knowlesi</i>
Emax_0331	44	9,00E-06	1,19E-02	10,0	<i>Similar to ACT_to XGO Actin</i>	AAC13766.1	<i>Toxoplasma gondii</i>

Tabela 9- *Contigs* diferencialmente expressos entre os estágios da esporulação de *E.maxima* que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), representado pela soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
Emax_0109	43	0,00E+00	1,21E-02	10,0	<i>Similar to hypothetical protein</i>	XP_001238682.1	<i>Eimeria tenella</i>
Emax_0955	43	0,00E+00	1,21E-02	10,0	<i>Similar to aminopeptidase n, putative</i>	XP_954343.1	<i>Theileria annulata</i>
Emax_0208	42	0,00E+00	1,23E-02	10,0	<i>Similar to microneme protein Etmic-2</i>	AAD05565.1	<i>Eimeria tenella</i>
Emax_0017	39	0,00E+00	1,31E-02	10,0	<i>Similar to heat shock protein 90</i>	AAS17969.1	<i>Eimeria acervulina</i>
Emax_0483	39	0,00E+00	1,31E-02	10,0	<i>Similar to hypothetical protein PFE0220w</i>	XP_001351603.1	<i>Plasmodium falciparum</i>
Emax_1088	35	4,35E-03	1,46E-02	7,0	<i>Similar to cysteine proteinase</i>	AAL60053.1	<i>Toxoplasma gondii</i>
Emax_0530	31	5,00E-06	1,64E-02	10,0	<i>Similar to hypothetical protein PF14_0571</i>	XP_001348745.1	<i>Plasmodium falciparum</i>
Emax_0090	30	1,00E-06	1,69E-02	10,0	<i>Similar to surface antigen 10</i>	ABR53732.1	<i>Eimeria tenella</i>
Emax_0287	30	8,00E-06	1,69E-02	10,0	<i>Similar to surface antigen 9</i>	CAE52300.1	<i>Eimeria tenella</i>
Emax_0325	28	0,00E+00	1,79E-02	10,0	<i>Similar to unknown</i>	ABQ41434.1	<i>Eimeria tenella</i>
Emax_1020	28	2,27E-03	1,79E-02	8,7	<i>Similar to actin depolymerizing facto r</i>	ABM89551.1	<i>Eimeria tenella</i>
Emax_1012	25	1,70E-02	1,96E-02	1,3	<i>Similar to surface antigen 4</i>	CAE52296.1	<i>Eimeria tenella</i>
Emax_0733	21	0,00E+00	2,22E-02	10,0	<i>Similar to transporter, putative</i>	XP_001613793.1	<i>Plasmodium vivax</i>
Emax_0011	20	0,00E+00	2,29E-02	10,0	<i>Similar to hypothetical protein PVX_115060</i>	XP_001616402.1	<i>Plasmodium vivax</i>
Emax_0054	20	5,40E-05	2,29E-02	10,0	<i>Similar to DNA-directed RNA polymerase II/III subunit</i>	XP_667260.1	<i>Cryptosporidium hominis</i>
Emax_0357	20	0,00E+00	2,29E-02	10,0	<i>Similar to DyNamin-like protein, putative</i>	CAJ20697.1	<i>Toxoplasma gondii</i>
Emax_0816	20	1,56E-02	2,29E-02	3,2	<i>Similar to eukaryotic translation initiation facto r 1</i>	ABB00912.1	<i>Toxoplasma gondii</i>
Emax_0960	20	5,40E-05	2,29E-02	10,0	<i>Similar to membrane skeleto n protein IMC2A</i>	AAK38356.3	<i>Toxoplasma gondii</i>

Tabela 9- *Contigs* diferencialmente expressos entre os estágios da esporulação de *E.maxima* que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), representado pela soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
Emax_0999	17	0,00E+00	2,52E-02	10,0	<i>Similar to rhomboid-like protein</i>	ABC50099.1	<i>Eimeria tenella</i>
Emax_0155	16	3,17E-03	2,60E-02	8,8	<i>Similar to hypothetical protein PY00232</i>	XP_730242.1	<i>Plasmodium yoelii</i>
Emax_0372	15	5,04E-04	2,68E-02	9,8	<i>Similar to structure specific recognition protein, putative</i>	XP_001611449.1	<i>Babesia bovis</i>
Emax_0735	15	0,00E+00	2,68E-02	10,0	<i>Similar to transporter, putative</i>	XP_001613793.1	<i>Plasmodium vivax</i>
Emax_1208	15	5,90E-03	2,68E-02	7,8	<i>Similar to small heat shock protein</i>	AAN87330.1	<i>Neospora caninum</i>
Emax_0222	14	1,44E-04	2,77E-02	9,9	<i>Similar to hypothetical protein, conserved</i>	EEA06007.1	<i>Cryptosporidium muris</i>
Emax_0243	14	3,30E-05	2,77E-02	10,0	<i>Similar to putative calcium binding protein precursor</i>	CAI72625.1	<i>Eimeria tenella</i>
Emax_0669	14	0,00E+00	2,77E-02	10,0	<i>Similar to hypothetical protein TA06810</i>	XP_954033.1	<i>Theileria annulata</i>
Emax_0990	14	1,08E-03	2,77E-02	9,6	<i>Similar to pre-mrna splicing factor protein, putative</i>	CAJ20726.1	<i>Toxoplasma gondii</i>
Emax_1228	14	0,00E+00	2,77E-02	10,0	<i>Similar to hypothetical protein</i>	XP_980864.1	<i>Mus musculus</i>
Emax_0002	13	0,00E+00	2,85E-02	10,0	<i>Similar to DEAD/DEAH box helicase</i>	XP_001609349.1	<i>Babesia bovis</i>
Emax_0273	13	5,31E-03	2,85E-02	8,1	<i>Similar to hypothetical protein LOC100124841</i>	NP_001096276.1	<i>Xenopus tropicalis</i>
Emax_0927	13	0,00E+00	2,85E-02	10,0	<i>Similar to 1433_EIMTE 14-3-3 protein</i>	AAD02687.1	<i>Eimeria tenella</i>
Emax_0115	12	2,80E-05	2,93E-02	10,0	<i>Similar to TBB_EIMTE Tubulin beta chain (Beta-tubulin)</i>	Q27380.1	<i>Eimeria tenella</i>
Emax_0292	12	5,00E-06	2,93E-02	10,0	<i>Similar to hypothetical protein</i>	XP_980864.1	<i>Mus musculus</i>
Emax_0833	12	2,91E-03	2,93E-02	9,0	<i>Similar to hypothetical protein TRIADDRAFT_26095</i>	XP_002112978.1	<i>Trichoplax adhaerens</i>
Emax_0912	12	2,91E-03	2,93E-02	9,0	<i>Similar to hypothetical protein, conserved in Plasmodium species</i>	CAQ42293.1	<i>Plasmodium knowlesi</i>
Emax_0479	11	5,58E-03	3,00E-02	8,1	<i>Similar to hypothetical protein CLOBOL_01718</i>	ZP_02084194.1	<i>Clostridium bolteae</i>

Tabela 9- *Contigs* diferencialmente expressos entre os estágios da esporulação de *E.maxima* que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), representado pela soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
E _{max} _0799	10	8,90E-03	3,07E-02	7,1	<i>Similar to DNA-directed RNA polymerase I subunit, putative</i>	CAQ38515.1	<i>Plasmodium knowlesi</i>
E _{max} _0815	10	2,40E-02	3,07E-02	2,2	<i>Similar to AF361949_1 histone 3</i>	AAM00267.1	<i>Eimeria tenella</i>
E _{max} _1006	10	8,90E-03	3,07E-02	7,1	<i>Similar to cell division cycle protein 48 homologue, putative</i>	CAQ40887.1	<i>Plasmodium knowlesi</i>
E _{max} _0051	9	1,31E-02	3,13E-02	5,8	<i>Similar to hypothetical protein, conserved</i>	XP_955121.1	<i>Theileria annulata</i>
E _{max} _0261	9	1,48E-02	3,13E-02	5,3	<i>Similar to B-box zinc finger domain-containing protein</i>	EEA08481.1	<i>Cryptosporidium muris</i>
E _{max} _0569	9	0,00E+00	3,13E-02	10,0	<i>Similar to RRPp/PMC2 like exosome 3'-5' exoribonuclease subunit with an RNaseD domain and an HRDc domain</i>	XP_625975.1	<i>Cryptosporidium parvum</i>
E _{max} _0610	9	1,31E-02	3,13E-02	5,8	<i>Similar to importin alpha</i>	XP_625580.1	<i>Cryptosporidium parvum</i>
E _{max} _0761	9	1,31E-02	3,13E-02	5,8	<i>Similar to hypothetical protein, conserved</i>	XP_001349941.1	<i>Plasmodium falciparum</i>
E _{max} _1225	9	0,00E+00	3,13E-02	10,0	<i>Similar to hypothetical protein</i>	XP_980864.1	<i>Mus musculus</i>
E _{max} _0276	8	2,50E-02	3,17E-02	2,1	<i>Similar to calmodulin</i>	XP_001348497.1	<i>Plasmodium falciparum</i>
E _{max} _0464	8	2,50E-02	3,17E-02	2,1	<i>Similar to actin-like protein 1</i>	AAW23163.1	<i>Toxoplasma gondii</i>
E _{max} _0507	8	2,50E-02	3,17E-02	2,1	<i>Similar to cholinephosphate cytidyltransferase</i>	XP_665462.1	<i>Cryptosporidium hominis</i>
E _{max} _0528	8	2,90E-05	3,17E-02	10,0	<i>Similar to aminopeptidase, putative</i>	XP_001610522.1	<i>Babesia bovis</i>
E _{max} _0531	8	2,50E-02	3,17E-02	2,1	<i>Similar to hypothetical protein, conserved</i>	XP_955400.1	<i>Theileria annulata</i>
E _{max} _0532	8	2,50E-02	3,17E-02	2,1	<i>Similar to multiprotein bridging factor type 1 like transcriptional co-activator</i>	XP_626922.1	<i>Cryptosporidium parvum</i>
E _{max} _0858	8	2,50E-02	3,17E-02	2,1	<i>Similar to mRNA cleavage factor-like protein, putative</i>	XP_001613540.1	<i>Plasmodium vivax</i>
E _{max} _0893	8	2,50E-02	3,17E-02	2,1	<i>Similar to ClpB</i>	YP_001498924.1	<i>Rickettsia massiliae</i>
E _{max} _1105	8	2,50E-02	3,17E-02	2,1	<i>Similar to bromodomain containing protein</i>	XP_001610355.1	<i>Babesia bovis</i>

Tabela 9- *Contigs* diferencialmente expressos entre os estágios da esporulação de *E.maxima* que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), representado pela soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
Emax_1179	8	2,50E-02	3,17E-02	2,1	<i>Similar to similar to retinoblasto ma binding protein 4 isoform 2</i>	XP_855677.1	<i>Canis familiaris</i>
Emax_1229	8	9,00E-06	3,17E-02	10,0	<i>Similar to hypoxanthine phosphoribosyltransferase</i>	XP_723630.1	<i>Plasmodium yoelii</i>
Emax_0329	7	3,90E-05	3,19E-02	10,0	<i>Similar to cystathione gamma lyase, putative</i>	XP_001568420.1	<i>Leishmania braziliensis</i>
Emax_0420	7	1,47E-03	3,19E-02	9,5	<i>Similar to hypothetical protein PY03786</i>	XP_723988.1	<i>Plasmodium yoelii</i>
Emax_0389	6	1,60E-04	3,16E-02	9,9	<i>Similar to unnamed protein product</i>	XP_451799.1	<i>Kluyveromyces lactis</i>
Emax_0708	6	1,60E-04	3,16E-02	9,9	<i>Similar to microneme protein MIC3</i>	AAR87667.1	<i>Eimeria tenella</i>
Emax_0793	6	7,57E-03	3,16E-02	7,6	<i>Similar to hypothetical protein, conserved</i>	EEA07916.1	<i>Cryptosporidium muris</i>
Emax_1001	6	1,60E-04	3,16E-02	9,9	<i>Similar to hypothetical protein PB300774.00.0</i>	XP_673909.1	<i>Plasmodium berghei</i>
Emax_1231	6	1,60E-04	3,16E-02	9,9	<i>Similar to CPD photo lyase</i>	BAE48230.1	<i>Chlorella pyrenoidosa</i>
Emax_0392	5	7,31E-04	3,08E-02	9,8	<i>Similar to hypothetical protein PY04760</i>	XP_725097.1	<i>Plasmodium yoelii</i>
Emax_0594	5	7,31E-04	3,08E-02	9,8	<i>Similar to heat shock protein 70</i>	AAP57537.3	<i>Locusta migratoria</i>
Emax_0666	5	7,31E-04	3,08E-02	9,8	<i>Similar to cell-cycle-related serine/threonine protein kinase, CDK homologue, putative</i>	XP_953424.1	<i>Theileria annulata</i>
Emax_1201	5	2,59E-03	3,08E-02	9,2	<i>Similar to putative rhoptry neck protein</i>	CAO79912.1	<i>Eimeria tenella</i>
Emax_0008	4	1,07E-02	2,89E-02	6,3	<i>Similar to aminopeptidase, putative</i>	XP_001610522.1	<i>Babesia bovis</i>
Emax_0012	4	1,67E-03	2,89E-02	9,4	<i>Similar to vacuolar protein sorting 29, putative</i>	CAQ42897.1	<i>Plasmodium knowlesi</i>
Emax_0039	4	1,67E-03	2,89E-02	9,4	<i>Similar to hypothetical protein, conserved in Apicomplexan species</i>	CAQ40676.1	<i>Plasmodium knowlesi</i>
Emax_0042	4	3,65E-03	2,89E-02	8,7	<i>Similar to PMCA-like calcium ATPase A1</i>	ABN55908.1	<i>Toxoplasma gondii</i>
Emax_0198	4	1,67E-03	2,89E-02	9,4	<i>Similar to predicted protein</i>	XP_001750709.1	<i>Monosiga brevicollis</i>

Tabela 9- *Contigs* diferencialmente expressos entre os estágios da esporulação de *E.maxima* que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), representado pela soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Conclusão)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
Emax_0595	4	1,82E-02	2,89E-02	3,7	<i>Similar to PP2C like protein phosphatase</i>	XP_625580.1	<i>Cryptosporidium parvum</i>
Emax_0706	4	3,65E-03	2,89E-02	8,7	<i>Similar to hypothetical protein PC000433.01.0</i>	XP_743776.1	<i>Plasmodium chabaudi</i>
Emax_0744	4	3,65E-03	2,89E-02	8,7	<i>Similar to membrane skeletal protein IMC1</i>	XP_001613793.1	<i>Toxoplasma gondii</i>
Emax_0861	4	2,85E-02	2,89E-02	0,1	<i>Similar to hypothetical protein</i>	XP_980864.1	<i>Mus musculus</i>
Emax_0969	4	3,65E-03	2,89E-02	8,7	<i>Similar to calcium-dependent protein kinase, putative</i>	CAJ20291.1	<i>Toxoplasma gondii</i>
Emax_1070	4	3,65E-03	2,89E-02	8,7	<i>Similar to ATP-binding cassette sub-family B member 2</i>	ACI24156.1	<i>Toxoplasma gondii</i>
Emax_0001	3	1,75E-02	2,55E-02	3,1	<i>Similar to clathrin assembly protein AP19</i>	XP_665272.1	<i>Cryptosporidium hominis</i>
Emax_0043	3	1,75E-02	2,55E-02	3,1	<i>Similar to similar to TATA box binding protein-related factor 2 CG18009-PD, isoform D isoform 1</i>	XP_396371.3	<i>Apis mellifera</i>
Emax_0045	3	1,75E-02	2,55E-02	3,1	<i>Similar to similar to cullin</i>	XP_967420.1	<i>Tribolium castaneum</i>
Emax_0172	3	1,75E-02	2,55E-02	3,1	<i>Similar to calcium-dependent protein kinase</i>	CAD32376.2	<i>Toxoplasma gondii</i>
Emax_0311	3	8,25E-03	2,55E-02	6,8	<i>Similar to Sis1-like protein</i>	ABY21519.1	<i>Toxoplasma gondii</i>
Emax_0416	3	1,75E-02	2,55E-02	3,1	<i>Similar to hypothetical protein PY00113</i>	XP_729062.1	<i>Plasmodium yoelii</i>
Emax_0438	3	8,25E-03	2,55E-02	6,8	<i>Similar to 2x WD domain containing protein</i>	XP_626548.1	<i>Cryptosporidium parvum</i>
Emax_0645	3	1,75E-02	2,55E-02	3,1	<i>Similar to orf344</i>	AAR90102.1	<i>Pseudomonas stutzeri</i>
Emax_0773	3	1,75E-02	2,55E-02	3,1	<i>Similar to TFP250</i>	AAO52676.1	<i>Eimeria maxima</i>
Emax_0977	3	1,75E-02	2,55E-02	3,1	<i>Similar to hypothetical protein TP02_0963</i>	XP_765376.1	<i>Theileria parva</i>
Emax_1000	3	1,75E-02	2,55E-02	3,1	<i>Similar to unknown</i>	ABQ41430.1	<i>Eimeria tenella</i>

Tabela 10- *Contigs* não diferencialmente expressos entre os estágios estudados que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), ou seja, soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continua)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
Emax_0268	2	2,03E-01	9,97E-03	0	<i>Similar to 40S ribosomal protein S16, putative</i>	EEA08558.1	<i>Cryptosporidium muris</i>
Emax_0698	2	1,00E+00	9,97E-03	0	<i>Similar to predicted protein</i>	XP_001637595.1	<i>Nematostella vectensis</i>
Emax_0792	2	4,12E-01	9,97E-03	0	<i>Similar to calmodulin-domain protein kinase</i>	CAA96439.1	<i>Eimeria tenella</i>
Emax_0904	2	1,00E+00	9,97E-03	0	<i>Similar to microneme protein MIC3</i>	AAR87666.1	<i>Eimeria tenella</i>
Emax_0005	2	8,47E-02	1,97E-02	0	<i>Similar to L-alanine dehydrogenase</i>	YP_678603.1	<i>Cytophaga hutchinsonii</i>
Emax_0006	2	8,47E-02	1,97E-02	0	<i>Similar to splicing factor, arginine/serine-rich 10 (transformer 2 homolog, Drosophila) variant</i>	BAD92445.1	<i>Homo sapiens</i>
Emax_0009	2	7,55E-01	1,97E-02	0	<i>Similar to A Chain A, Crystal Structure Of toxoplasma Gondii Peptidyl-Prolyl Cis- Trans Isomerase, 541.M00136</i>	pdb 3BO7	<i>Toxoplasma gondii</i>
Emax_0049	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein Chro.40217</i>	XP_667870.1	<i>Cryptosporidium hominis</i>
Emax_0068	2	5,15E-01	1,97E-02	0	<i>Similar to calcium-dependent protein kinase</i>	CAD32376.2	<i>Toxoplasma gondii</i>
Emax_0074	2	5,15E-01	1,97E-02	0	<i>Similar to AF288603_1 cAMP-dependent protein kinase regulatory subunit</i>	AAK01548.1	<i>Toxoplasma gondii</i>
Emax_0112	2	1,69E-01	1,97E-02	0	<i>Similar to glucose-6-phosphate dehydrogenase</i>	ABF20372.1	<i>Leishmania gerbilli</i>
Emax_0114	2	4,15E-02	1,97E-02	0	<i>Similar to Pfs77, putative</i>	XP_001616226.1	<i>Plasmodium vivax</i>
Emax_0121	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein, conserved</i>	XP_954760.1	<i>Theileria annulata</i>
Emax_0128	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein, conserved in Apicomplexan species</i>	CAQ42381.1	<i>Plasmodium knowlesi</i>
Emax_0129	2	5,15E-01	1,97E-02	0	<i>Similar to protein tyrosine phosphatase, putative</i>	CAJ20581.1	<i>Toxoplasma gondii</i>
Emax_0132	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein Bcep1808_1630</i>	YP_001119472.1	<i>Burkholderia vietnamiensis</i>
Emax_0143	2	8,47E-02	1,97E-02	0	<i>Similar to serine hydroxymethyltransferase</i>	AAT74582.1	<i>Toxoplasma gondii</i>
Emax_0146	2	8,47E-02	1,97E-02	0	<i>Similar to similar to RE14259p</i>	XP_001602628.1	<i>Nasonia vitripennis</i>

Tabela 10- *Contigs* não diferencialmente expressos entre os estágios estudados que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE). ou seja, soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
Emax_0149	2	1,00E+00	1,97E-02	0	<i>Similar to inner membrane complex associated protein 5</i>	ABI34611.1	<i>Toxoplasma gondii</i>
Emax_0165	2	5,15E-01	1,97E-02	0	<i>Similar to similar to Dynein heavy chain 5, axonemal (Axonemal beta dynein heavy chain 5) (Ciliary dynein heavy chain 5)</i>	XP_002122887.1	<i>Ciona intestinalis</i>
Emax_0179	2	5,15E-01	1,97E-02	0	<i>Similar to eukaryotic initiation facto r</i>	XP_001693606.1	<i>Chlamydomonas reinhardtii</i>
Emax_0207	2	5,15E-01	1,97E-02	0	<i>Similar to sporozoite surface protein</i>	AAS66744.1	<i>Eimeria acervulina</i>
Emax_0234	2	5,15E-01	1,97E-02	0	<i>Similar to perforin-like protein 1</i>	ABK97634.2	<i>Toxoplasma gondii</i>
Emax_0244	2	7,55E-01	1,97E-02	0	<i>Similar to AF465543_1 cGMP-dependent protein kinase</i>	AAM22643.1	<i>Eimeria maxima</i>
Emax_0250	2	5,15E-01	1,97E-02	0	<i>Similar to glycosyltransferase 28 domain containing 1 (ISS)</i>	CAL55994.1	<i>Ostreococcus tauri</i>
Emax_0254	2	5,15E-01	1,97E-02	0	<i>Similar to heat shock protein</i>	CAA81135.1	<i>Eimeria acervulina</i>
Emax_0282	2	5,15E-01	1,97E-02	0	<i>Similar to AF333959_1 calmodulin-domain protein kinase 2</i>	AAG53994.1	<i>Toxoplasma gondii</i>
Emax_0289	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein PVX_117635</i>	XP_001615821.1	<i>Plasmodium vivax</i>
Emax_0305	2	5,15E-01	1,97E-02	0	<i>Similar to predicted protein</i>	XP_001754931.1	<i>Physcomitrella patens</i>
Emax_0319	2	4,15E-02	1,97E-02	0	<i>Similar to glucose 6-phosphate-1-dehydrogenase</i>	XP_644436.1	<i>Dictyostelium discoideum</i>
Emax_0334	2	5,15E-01	1,97E-02	0	<i>Similar to adapto r-related protein complex 3, sigma 2 subunit, putative</i>	XP_001614031.1	<i>Plasmodium vivax</i>
Emax_0344	2	5,15E-01	1,97E-02	0	<i>Similar to protein serine/threonine phosphatase, putative</i>	CAQ41472.1	<i>Plasmodium knowlesi</i>
Emax_0366	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein BBOV_I004350</i>	XP_001609084.1	<i>Babesia bovis</i>
Emax_0384	2	1,00E+00	1,97E-02	0	<i>Similar to membrane anchor for myosin XIV precursor</i>	AAT07037.1	<i>Toxoplasma gondii</i>
Emax_0385	2	8,47E-02	1,97E-02	0	<i>Similar to RL5_EIMTE 60S ribosomal protein L5</i>	AAT97351.1	<i>Eimeria tenella</i>

Tabela 10- *Contigs* não diferencialmente expressos entre os estágios estudados que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), ou seja, soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
Emax_0390	2	8,47E-02	1,97E-02	0	<i>Similar to unknown</i>	ABQ41427.1	<i>Eimeria tenella</i>
Emax_0393	2	8,47E-02	1,97E-02	0	<i>Similar to similar to Inosito 1,4,5,-tris-phosphate receptor CG1063-PB, isoform B</i>	XP_392236.3	<i>Apis mellifera</i>
Emax_0405	2	5,15E-01	1,97E-02	0	<i>Similar to sec63 domain containing protein</i>	XP_001611230.1	<i>Babesia bovis</i>
Emax_0442	2	4,15E-02	1,97E-02	0	<i>Similar to endonuclease/exonuclease/phosphatase family protein</i>	EEA06268.1	<i>Cryptosporidium muris</i>
Emax_0443	2	4,15E-02	1,97E-02	0	<i>Similar to hypothetical protein BBOV_IV008800</i>	XP_001610802.1	<i>Babesia bovis</i>
Emax_0446	2	5,15E-01	1,97E-02	0	<i>Similar to EMB2765 (EMBRYO DEFECTIVE 2765)</i>	AAM20715.1	<i>Arabidopsis thaliana</i>
Emax_0458	2	5,15E-01	1,97E-02	0	<i>Similar to inosito 1 phosphatase, putative</i>	CAQ37870.1	<i>Plasmodium knowlesi</i>
Emax_0476	2	5,15E-01	1,97E-02	0	<i>Similar to em100 gene is homologous the Eimeria tenella gene et100 (accession number M73495) encoding the microneme protein Etp100</i>	AAA29076.1	<i>Eimeria tenella</i>
Emax_0514	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein, conserved</i>	XP_955533.1	<i>Theileria annulata</i>
Emax_0545	2	1,69E-01	1,97E-02	0	<i>Similar to predicted protein</i>	XP_001633749.1	<i>Nematostella vectensis</i>
Emax_0562	2	8,47E-02	1,97E-02	0	<i>Similar to hesB-like domain-containing protein</i>	AAT57939.1	<i>Toxoplasma gondii</i>
Emax_0580	2	5,15E-01	1,97E-02	0	<i>Similar to similar to ENSANGP00000027572</i>	XP_001603780.1	<i>Nasonia vitripennis</i>
Emax_0587	2	5,15E-01	1,97E-02	0	<i>Similar to xylose transport permease</i>	AAB46620.1	<i>Bacillus sp.</i>
Emax_0600	2	4,15E-02	1,97E-02	0	<i>Similar to Alkyl-Dihydroxyaceto nephosphate Synthase family member (ads-1)</i>	NP_497185.1	<i>Caenorhabditis elegans</i>
Emax_0607	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein PFL0650c</i>	XP_001350539.1	<i>Plasmodium falciparum</i>
Emax_0622	2	5,15E-01	1,97E-02	0	<i>Similar to Signal peptide peptidase, putative</i>	CAQ41568.1	<i>Plasmodium knowlesi</i>
Emax_0630	2	5,15E-01	1,97E-02	0	<i>Similar to transcription initiation factor TFIIB</i>	XP_764767.1	<i>Theileria parva</i>

Tabela 10- *Contigs* não diferencialmente expressos entre os estágios estudados que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), ou seja, soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
Emax_0633	2	5,15E-01	1,97E-02	0	<i>Similar to heat shock protein 70</i>	AAV63878.2	<i>Cyclospora cercopitheci</i>
Emax_0649	2	8,47E-02	1,97E-02	0	<i>Similar to hypothetical protein PF10_0186</i>	XP_001347471.1	<i>Plasmodium falciparum</i>
Emax_0653	2	1,00E+00	1,97E-02	0	<i>Similar to phosphatidylinosito 1-4-phosphate 5-kinase, putative</i>	XP_001388435.1	<i>Cryptosporidium parvum</i>
Emax_0684	2	4,15E-02	1,97E-02	0	<i>Similar to hypothetical protein PVX_123605</i>	XP_001617283.1	<i>Plasmodium vivax</i>
Emax_0690	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein with 5 transmembrane domains</i>	XP_626259.1	<i>Cryptosporidium parvum</i>
Emax_0691	2	5,15E-01	1,97E-02	0	<i>Similar to similar to DEAD (Asp-Glu-Ala-Asp) box polypeptide 52</i>	XP_001607535.1	<i>Nasonia vitripennis</i>
Emax_0694	2	5,15E-01	1,97E-02	0	<i>Similar to periplasmic binding protein</i>	ZP_01916543.1	<i>Limnobacter sp.</i>
Emax_0700	2	8,47E-02	1,97E-02	0	<i>Similar to choline-phosphate cytidyltransferase protein, putative</i>	EEA05501.1	<i>Cryptosporidium muris</i>
Emax_0713	2	5,15E-01	1,97E-02	0	<i>Similar to protein kinase 3</i>	AAC27446.1	<i>Toxoplasma gondii</i>
Emax_0714	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein</i>	CAN65823.1	<i>Vitis vinifera</i>
Emax_0721	2	1,69E-01	1,97E-02	0	<i>Similar to alkylglycerone phosphate synthase</i>	NP_001086573.1	<i>Xenopus laevis</i>
Emax_0722	2	1,69E-01	1,97E-02	0	<i>Similar to mito gen-activated protein kinase, putative</i>	CAJ20514.1	<i>Toxoplasma gondii</i>
Emax_0724	2	1,69E-01	1,97E-02	0	<i>Similar to Os02g0280100</i>	NP_001046553.1	<i>Oryza sativa</i>
Emax_0727	2	4,15E-02	1,97E-02	0	<i>Similar to predicted protein</i>	XP_001781353.1	<i>Physcomitrella patens</i>
Emax_0732	2	5,15E-01	1,97E-02	0	<i>Similar to unnamed protein product</i>	CAG01206.1	<i>Tetraodon nigroviridis</i>
Emax_0746	2	8,47E-02	1,97E-02	0	<i>Similar to cyclophilin</i>	XP_001347841.1	<i>Plasmodium falciparum</i>
Emax_0766	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein, conserved</i>	XP_954751.1	<i>Theileria annulata</i>
Emax_0797	2	7,55E-01	1,97E-02	0	<i>Similar to RIO1 family protein</i>	XP_001612063.1	<i>Babesia bovis</i>
Emax_0800	2	5,15E-01	1,97E-02	0	<i>Similar to 2-oxoglutarate dehydrogenase E1 component</i>	XP_001614954.1	<i>Plasmodium vivax</i>
Emax_0813	2	4,15E-02	1,97E-02	0	<i>Similar to folate/biopterin transporter family protein</i>	EEA06295.1	<i>Cryptosporidium muris</i>

Tabela 10- *Contigs* não diferencialmente expressos entre os estágios estudados que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), ou seja, soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
Emax_0814	2	4,15E-02	1,97E-02	0	<i>Similar to ATP-binding cassette protein subfamily B member 2</i>	ABN95811.1	<i>Toxoplasma gondii</i>
Emax_0825	2	5,15E-01	1,97E-02	0	<i>Similar to Ras family protein</i>	EEA04937.1	<i>Cryptosporidium muris</i>
Emax_0844	2	5,15E-01	1,97E-02	0	<i>Similar to putative transmembrane protein</i>	XP_642710.1	<i>Dictyostelium discoideum</i>
Emax_0863	2	1,69E-01	1,97E-02	0	<i>Similar to 2-methylcitrate synthase</i>	ABC59525.1	<i>Toxoplasma gondii</i>
Emax_0865	2	5,15E-01	1,97E-02	0	<i>Similar to similar to AGAP008746-PA</i>	XP_971591.1	<i>Tribolium castaneum</i>
Emax_0867	2	5,15E-01	1,97E-02	0	<i>Similar to 26S proteasome regulatory subunit S10b like AAA+ ATPase</i>	XP_626993.1	<i>Cryptosporidium parvum</i>
Emax_0871	2	5,15E-01	1,97E-02	0	<i>Similar to splicing factor 3a protein, putative</i>	XP_001608903.1	<i>Babesia bovis</i>
Emax_0879	2	1,00E+00	1,97E-02	0	<i>Similar to Thioredoxin family protein</i>	XP_001610156.1	<i>Babesia bovis</i>
Emax_0897	2	5,15E-01	1,97E-02	0	<i>Similar to surface antigen 12</i>	CAE52295.1	<i>Eimeria tenella</i>
Emax_0941	2	5,15E-01	1,97E-02	0	<i>Similar to fructose-1,6-bisphosphate aldolase</i>	AAN75043.1	<i>Toxoplasma gondii</i>
Emax_0957	2	5,15E-01	1,97E-02	0	<i>Similar to putative 50S ribosomal subunit L17</i>	AAT41970.1	<i>Fremyella diplosiphon</i>
Emax_0966	2	8,47E-02	1,97E-02	0	<i>Similar to hypothetical protein Chro.30352</i>	XP_665995.1	<i>Cryptosporidium hominis</i>
Emax_0968	2	8,47E-02	1,97E-02	0	<i>Similar to hypothetical protein, conserved in Plasmodium species</i>	CAQ41335.1	<i>Plasmodium knowlesi</i>
Emax_0975	2	8,47E-02	1,97E-02	0	<i>Similar to similar to GlcT-1 CG6437-PA</i>	XP_395023.3	<i>Apis mellifera</i>
Emax_0988	2	5,15E-01	1,97E-02	0	<i>Similar to similar to CG8877-PA</i>	XP_624014.2	<i>Apis mellifera</i>
Emax_0998	2	8,47E-02	1,97E-02	0	<i>Similar to hypothetical protein MAL13P1.308</i>	XP_001350329.1	<i>Plasmodium falciparum</i>
Emax_1013	2	5,15E-01	1,97E-02	0	<i>Similar to rhoptry protein 18</i>	CAJ27113.1	<i>Toxoplasma gondii</i>
Emax_1022	2	4,15E-02	1,97E-02	0	<i>Similar to hypothetical protein TA20230</i>	XP_954205.1	<i>Theileria annulata</i>
Emax_1027	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein, conserved</i>	XP_001348873.1	<i>Plasmodium falciparum</i>

Tabela 10- *Contigs* não diferencialmente expressos entre os estágios estudados que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), ou seja, soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
E _{max} _1043	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein PC302260.00.0</i>	XP_740252.1	<i>Plasmodium chabaudi</i>
E _{max} _1044	2	5,15E-01	1,97E-02	0	<i>Similar to heat-shock protein</i>	CAA87085.1	<i>Eimeria maxima</i>
E _{max} _1048	2	5,15E-01	1,97E-02	0	<i>Similar to GE18776</i>	XP_002088822.1	<i>Drosophila yakuba</i>
E _{max} _1059	2	1,69E-01	1,97E-02	0	<i>Similar to formin-related protein</i>	XP_667351.1	<i>Cryptosporidium hominis</i>
E _{max} _1061	2	8,47E-02	1,97E-02	0	<i>Similar to hypothetical protein, expressed</i>	CAJ20390.1	<i>Toxoplasma gondii</i>
E _{max} _1065	2	5,15E-01	1,97E-02	0	<i>Similar to phosphatidylserine decarboxylase</i>	ABE03002.1	<i>Toxoplasma gondii</i>
E _{max} _1067	2	1,69E-01	1,97E-02	0	<i>Similar to alpha tubulin</i>	CAA61255.1	<i>Eimeria acervulina</i>
E _{max} _1068	2	8,47E-02	1,97E-02	0	<i>Similar to Geranylgeranyl transferase type-2 subunit beta</i>	ACI33626.1	<i>Salmo salar</i>
E _{max} _1084	2	5,15E-01	1,97E-02	0	<i>Similar to conserved hypothetical protein</i>	XP_625894.1	<i>Cryptosporidium parvum</i>
E _{max} _1093	2	5,15E-01	1,97E-02	0	<i>Similar to 26S protease regulatory subunit 8, putative</i>	XP_001614014.1	<i>Plasmodium vivax</i>
E _{max} _1112	2	5,15E-01	1,97E-02	0	<i>Similar to 60S ribosomal protein I3, putative</i>	XP_954213.1	<i>Theileria annulata</i>
E _{max} _1135	2	7,55E-01	1,97E-02	0	<i>Similar to predicted protein</i>	XP_001748524.1	<i>Monosiga brevicollis</i>
E _{max} _1145	2	8,47E-02	1,97E-02	0	<i>Similar to hypothetical protein PB301386.00.0</i>	XP_672836.1	<i>Plasmodium berghei</i>
E _{max} _1146	2	8,47E-02	1,97E-02	0	<i>Similar to replication factor A protein</i>	XP_765009.1	<i>Theileria parva</i>
E _{max} _1150	2	1,00E+00	1,97E-02	0	<i>Similar to translation elongation factor EF-1, putative</i>	CAQ40009.1	<i>Plasmodium knowlesi</i>
E _{max} _1160	2	5,15E-01	1,97E-02	0	<i>Similar to molecular chaperone, putative</i>	XP_954074.1	<i>Theileria annulata</i>
E _{max} _1169	2	5,15E-01	1,97E-02	0	<i>Similar to hypothetical protein TRIADDRAFT_22483</i>	XP_002111206.1	<i>Trichoplax adhaerens</i>
E _{max} _1182	2	5,15E-01	1,97E-02	0	<i>Similar to predicted protein</i>	XP_001771583.1	<i>Physcomitrella patens</i>
E _{max} _1194	2	4,15E-02	1,97E-02	0	<i>Similar to hypothetical protein, conserved in Plasmodium species</i>	CAQ39612.1	<i>Plasmodium knowlesi</i>

Tabela 10- *Contigs* não diferencialmente expressos entre os estágios estudados que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), ou seja, soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	score	Resultado da anotação automática	Código de acesso	Organismo
E _{max} _1204	2	8,47E-02	1,97E-02	0	<i>Similar to XPG I-region family protein</i>	XP_001609909.1	<i>Babesia bovis</i>
E _{max} _1215	2	5,15E-01	1,97E-02	0	<i>Similar to D-isomer specific 2-hydroxyacid dehydrogenase, NAD-binding</i>	YP_001174908.1	<i>Enterobacter sp.</i>
E _{max} _1224	2	1,00E+00	1,97E-02	0	<i>Similar to PGPD14 protein with at least one predicted RING finger, possible plant origin</i>	XP_628085.1	<i>Cryptosporidium parvum</i>
E _{max} _1226	2	1,69E-01	1,97E-02	0	<i>Similar to isoleucyl-tRNA syntehtase</i>	XP_728219.1	<i>Plasmodium yoelii</i>
E _{max} _0058	3	4,23E-01	2,55E-02	0	<i>Similar to hypothetical protein BBOV_I1003090</i>	XP_001609832.1	<i>Babesia bovis</i>
E _{max} _0077	3	4,23E-01	2,55E-02	0	<i>Similar to GNAT family histone acetyltransferase GCN5-B</i>	AAW72884.1	<i>Toxoplasma gondii</i>
E _{max} _0102	3	8,50E-01	2,55E-02	0	<i>Similar to MYOA_to XGO RecName: Full=Myosin-A; Short=MyoA; AltName: Full=TgM-A</i>	AAC47724.1	<i>Toxoplasma gondii</i>
E _{max} _0169	3	6,34E-01	2,55E-02	0	<i>Similar to predicted protein</i>	XP_001764982.1	<i>Physcomitrella patens</i>
E _{max} _0178	3	4,23E-01	2,55E-02	0	<i>Similar to iswi protein</i>	XP_676569.1	<i>Plasmodium berghei</i>
E _{max} _0255	3	4,23E-01	2,55E-02	0	<i>Similar to signal recognition particle protein SRP54, putative</i>	EEA06981.1	<i>Cryptosporidium muris</i>
E _{max} _0288	3	4,23E-01	2,55E-02	0	<i>Similar to predicted protein</i>	XP_001753311.1	<i>Physcomitrella patens</i>
E _{max} _0295	3	4,31E-02	2,55E-02	0	<i>Similar to phosphatidylinositol N-acetylglucosaminyltransferase subunit C</i>	CAI84649.1	<i>Toxoplasma gondii</i>
E _{max} _0371	3	4,23E-01	2,55E-02	0	<i>Similar to insulysin, putative</i>	EEA08436.1	<i>Cryptosporidium muris</i>
E _{max} _0463	3	4,23E-01	2,55E-02	0	<i>Similar to ATP-dependent protease Clp ATP-binding subunit</i>	YP_001692820.1	<i>Finexgoldia magna</i>
E _{max} _0468	3	4,23E-01	2,55E-02	0	<i>Similar to hypothetical protein</i>	CAJ20703.1	<i>Toxoplasma gondii</i>
E _{max} _0487	3	4,23E-01	2,55E-02	0	<i>Similar to UBA/Ts-N domain-containing protein</i>	EEA07426.1	<i>Cryptosporidium muris</i>
E _{max} _0490	3	2,19E-01	2,55E-02	0	<i>Similar to membrane skeletal protein IMC1</i>	AAK39634.	<i>Toxoplasma gondii</i>
E _{max} _0508	3	4,23E-01	2,55E-02	0	<i>Similar to 6-phosphogluconate dehydrogenase NAD-binding</i>	ZP_02912884.1	<i>Geobacillus sp</i>

Tabela 10- *Contigs* não diferencialmente expressos entre os estágios estudados que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), ou seja, soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
Emax_0510	3	4,23E-01	2,55E-02	0	<i>Similar to hypothetical protein OsJ_013798</i>	EAZ30315.1	<i>Oryza sativa</i>
Emax_0548	3	1,00E+00	2,55E-02	0	<i>Similar to zincin/aminopeptidase N like metalloprotease</i>	XP_627236.1	<i>Cryptosporidium parvum</i>
Emax_0574	3	4,23E-01	2,55E-02	0	<i>Similar to hypothetical protein, expressed</i>	CAJ20390.1	<i>Toxoplasma gondii</i>
Emax_0579	3	4,23E-01	2,55E-02	0	<i>Similar to AF518411_1 facilitative glucose transporter; GT1</i>	AAM69350.1	<i>Toxoplasma gondii</i>
Emax_0588	3	4,23E-01	2,55E-02	0	<i>Similar to hypothetical protein GSPATT00032284001</i>	XP_001429450.1	<i>Paramecium tetraurelia</i>
Emax_0597	3	7,01E-02	2,55E-02	0	<i>Similar to ribosomal protein L3, putative</i>	XP_001608935.1	<i>Babesia bovis</i>
Emax_0635	3	4,23E-01	2,55E-02	0	<i>Similar to membrane acyl-CoA binding protein</i>	AAT81164.1	<i>Agave americana</i>
Emax_0655	3	4,23E-01	2,55E-02	0	<i>Similar to protein phosphatase 2C2</i>	ABV44288.1	<i>Toxoplasma gondii</i>
Emax_0696	3	4,23E-01	2,55E-02	0	<i>Similar to hypothetical protein, conserved</i>	XP_001349096.1	<i>Plasmodium falciparum</i>
Emax_0731	3	4,23E-01	2,55E-02	0	<i>Similar to hypothetical protein TA02855</i>	XP_954113.1	<i>Theileria annulata</i>
Emax_0781	3	2,19E-01	2,55E-02	0	<i>Similar to protein phosphatase, putative</i>	CAJ20491.1	<i>Toxoplasma gondii</i>
Emax_0791	3	4,23E-01	2,55E-02	0	<i>Similar to nucleosome assembly protein 19075338, putative</i>	XP_001238725.1	<i>Eimeria tenella</i>
Emax_0823	3	1,00E+00	2,55E-02	0	<i>Similar to proliferation-associated protein 2g4</i>	XP_765891.1	<i>Theileria parva</i>
Emax_0848	3	1,43E-01	2,55E-02	0	<i>Similar to hypothetical protein CC1G_01399</i>	XP_001837487.1	<i>Coprinopsis cinerea</i>
Emax_0891	3	4,23E-01	2,55E-02	0	<i>Similar to hypothetical protein TP02_0067</i>	XP_764636.1	<i>Theileria parva</i>
Emax_0924	3	4,23E-01	2,55E-02	0	<i>Similar to protein disulfide-isomerase domain-containing protein</i>	EEA06745.1	<i>Cryptosporidium muris</i>
Emax_1009	3	4,23E-01	2,55E-02	0	<i>Similar to hypothetical protein cgd4_3330</i>	XP_001388076.1	<i>Cryptosporidium parvum</i>
Emax_1017	3	4,23E-01	2,55E-02	0	<i>Similar to inner membrane complex associated protein 5</i>	ABI34611.1	<i>Toxoplasma gondii</i>
Emax_1035	3	8,50E-01	2,55E-02	0	<i>Similar to hypothetical protein MAL8P1.30</i>	XP_001349277.1	<i>Plasmodium falciparum</i>
Emax_1045	3	4,23E-01	2,55E-02	0	<i>Similar to hypothetical protein PC000096.02.0</i>	XP_741463.1	<i>Plasmodium chabaudi</i>

Tabela 10- *Contigs* não diferencialmente expressos entre os estágios estudados que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), ou seja, soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
E _{max} _1054	3	6,92E-02	2,55E-02	0	<i>Similar to acyltransferase, putative</i>	CAQ41697.1	<i>Plasmodium knowlesi</i>
E _{max} _1085	3	4,23E-01	2,55E-02	0	<i>Similar to cAMP-dependent protein kinase catalytic subunit isoform 2</i>	AAX55640.1	<i>Toxoplasma gondii</i>
E _{max} _1102	3	4,23E-01	2,55E-02	0	<i>Similar to poly(A) polymerase protein, putative</i>	EEA06148.1	<i>Cryptosporidium muris</i>
E _{max} _1180	3	4,23E-01	2,55E-02	0	<i>Similar to hypothetical protein GSPATT00001645001</i>	XP_001446001.1	<i>Paramecium tetraurelia</i>
E _{max} _1184	3	4,23E-01	2,55E-02	0	<i>Similar to Plasmodium falciparum asparagine-rich protein</i>	XP_001349865.1	<i>Plasmodium falciparum</i>
E _{max} _1192	3	4,31E-02	2,55E-02	0	<i>Similar to glycine-rich nucleic acid binding protein</i>	ACF24575.1	<i>Gymnochlora stellata</i>
E _{max} _1213	3	4,23E-01	2,55E-02	0	<i>Similar to zinc finger protein</i>	XP_626343.1	<i>Cryptosporidium parvum</i>
E _{max} _1220	3	6,34E-01	2,55E-02	0	<i>Similar to DNAJ heat shock N-terminal domain-containing protein / cell division protein-related</i>	AAG51437.1	<i>Arabidopsis thaliana</i>
E _{max} _1227	3	1,43E-01	2,55E-02	0	<i>Similar to hypothetical protein</i>	XP_980864.1	<i>Mus musculus</i>
E _{max} _0014	4	1,89E-01	2,89E-02	0	<i>Similar to hypothetical protein COPEUT_01588</i>	ZP_02206798.1	<i>Coprococcus eutactus</i>
E _{max} _0126	4	8,25E-01	2,89E-02	0	<i>Similar to hypothetical protein PVX_122640</i>	XP_001617091.1	<i>Plasmodium vivax</i>
E _{max} _0205	4	1,89E-01	2,89E-02	0	<i>Similar to Os01g0888700</i>	NP_001045041.1	<i>Oryza sativa</i>
E _{max} _0341	4	1,89E-01	2,89E-02	0	<i>Similar to 5'-3' exoribonuclease 1 (ISS)</i>	CAL53622.1	<i>Ostreococcus tauri</i>
E _{max} _0401	4	1,89E-01	2,89E-02	0	<i>Similar to organellar heat shock protein</i>	CAA87086.1	<i>Eimeria tenella</i>
E _{max} _0423	4	8,25E-01	2,89E-02	0	<i>Similar to CDPK2_PLAFK Calcium-dependent protein kinase 2 (PfCDPK2)</i>	CAA68090.1	<i>Plasmodium falciparum</i>
E _{max} _0522	4	4,82E-02	2,89E-02	0	<i>Similar to PMCA-like calcium ATPase A1</i>	ABN55908.1	<i>Toxoplasma gondii</i>
E _{max} _0583	4	1,89E-01	2,89E-02	0	<i>Similar to dynein intermediate chain 2, putative</i>	CAJ20630.1	<i>Toxoplasma gondii</i>
E _{max} _0604	4	4,89E-01	2,89E-02	0	<i>Similar to ubiquitin-conjugating enzyme family protein</i>	EEA07656.1	<i>Cryptosporidium muris</i>

Tabela 10- *Contigs* não diferencialmente expressos entre os estágios estudados que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), ou seja, soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
E _{max} _0624	4	1,89E-01	2,89E-02	0	<i>Similar to transcription termination factor Rho</i>	YP_001538809.1	<i>Salinispora arenicola</i>
E _{max} _0736	4	6,55E-01	2,89E-02	0	<i>Similar to hypothetical protein PF14_0155</i>	XP_001348328.1	<i>Plasmodium falciparum</i>
E _{max} _0901	4	3,66E-01	2,89E-02	0	<i>Similar to hypothetical protein PC000096.00.0</i>	XP_745589.1	<i>Plasmodium chabaudi</i>
E _{max} _0939	4	1,89E-01	2,89E-02	0	<i>Similar to helicase with zinc finger motif protein, putative</i>	XP_001019542.2	<i>Tetrahymena thermophila</i>
E _{max} _1122	4	1,89E-01	2,89E-02	0	<i>Similar to conserved expressed protein</i>	XP_626656.1	<i>Cryptosporidium parvum</i>
E _{max} _1186	4	1,89E-01	2,89E-02	0	<i>Similar to phosphate transporter</i>	CAE30463.1	<i>Plasmodium falciparum</i>
E _{max} _0056	5	9,28E-02	3,08E-02	0	<i>Similar to hypothetical protein PVX_083230</i>	XP_001614245.1	<i>Plasmodium vivax</i>
E _{max} _0224	5	9,28E-02	3,08E-02	0	<i>Similar to Ubiquitin-like protein, putative</i>	CAQ40535.1	<i>Plasmodium knowlesi</i>
E _{max} _0230	5	9,28E-02	3,08E-02	0	<i>Similar to acidic ribosomal protein P2</i>	AAL50112.1	<i>Cyclospora cayetanensis</i>
E _{max} _0235	5	1,96E-01	3,08E-02	0	<i>Similar to putative peptidyl-prolyl cis-trans isomerase</i>	YP_001827462.1	<i>Streptomyces griseus</i>
E _{max} _0353	5	9,28E-02	3,08E-02	0	<i>Similar to TATA element modulatory factor</i>	XP_729580.1	<i>Plasmodium yoelii</i>
E _{max} _0373	5	1,22E-01	3,08E-02	0	<i>Similar to RHBL2_to XGO Rhomboid-like protease 2</i>	AAT29066.1	<i>Toxoplasma gondii</i>
E _{max} _0425	5	9,28E-02	3,08E-02	0	<i>Similar to ATP-dependent RNA helicase</i>	XP_763560.1	<i>Theileria parva</i>
E _{max} _0469	5	9,28E-02	3,08E-02	0	<i>Similar to myosin light chain TgMLC1</i>	AAL08211.1	<i>Toxoplasma gondii</i>
E _{max} _0623	5	9,28E-02	3,08E-02	0	<i>Similar to glutaredoxin-like protein</i>	XP_765847.1	<i>Theileria parva</i>
E _{max} _0757	5	9,28E-02	3,08E-02	0	<i>Similar to AF411961_1 cGMP-dependent protein kinase</i>	AAM20900.1	<i>Eimeria tenella</i>
E _{max} _0930	5	9,28E-02	3,08E-02	0	<i>Similar to akt protein</i>	ACH73234.1	<i>Saccoglossus kowalevskii</i>
E _{max} _0987	5	9,28E-02	3,08E-02	0	<i>Similar to similar to transaldolase</i>	XP_001602166.1	<i>Nasonia vitripennis</i>
E _{max} _1063	5	9,28E-02	3,08E-02	0	<i>Similar to ATP-dependent chaperone ClpB</i>	YP_001560905.1	<i>Clostridium phytofermentans</i>
E _{max} _1141	5	9,28E-02	3,08E-02	0	<i>Similar to Hsp70 interacting protein</i>	ABG75606.1	<i>Toxoplasma gondii</i>

Tabela 10- *Contigs* não diferencialmente expressos entre os estágios estudados que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), ou seja, soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
Emax_1177	5	9,28E-02	3,08E-02	0	<i>Similar to iron regulato ry protein-like protein</i>	AAT68238.1	<i>Toxoplasma gondii</i>
Emax_1218	5	9,28E-02	3,08E-02	0	<i>Similar to SERPIN1 protein precursor</i>	CAI72624.1	<i>Eimeria tenella</i>
Emax_0167	6	8,10E-02	3,16E-02	0	<i>Similar to calcium/calmodulin dependent protein kinase with a kinase domain and 4 calmodulin like EF hands</i>	XP_626355.1	<i>Cryptosporidium parvum</i>
Emax_0431	6	2,63E-01	3,16E-02	0	<i>Similar to heat shock protein 90</i>	AAS18319.1	<i>Eimeria acervulina</i>
Emax_0765	6	8,10E-02	3,16E-02	0	<i>Similar to 2-oxoglutarate dehydrogenase, E1 component</i>	XP_730299.1	<i>Plasmodium yoelii</i>
Emax_0769	6	8,10E-02	3,16E-02	0	<i>Similar to TFP250</i>	AAO52676.1	<i>Eimeria maxima</i>
Emax_1002	6	8,10E-02	3,16E-02	0	<i>Similar to CALM_ALEFU Calmodulin (CaM)</i>	A4UHC0.1	<i>Pfiesteria piscicida</i>
Emax_1037	6	9,39E-02	3,16E-02	0	<i>Similar to hypothetical protein PC000755.04.0</i>	XP_744532.1	<i>Plasmodium chabaudi</i>
Emax_0284	7	3,61E-02	3,19E-02	0	<i>Similar to hypothetical protein PY04677</i>	XP_725006.1	<i>Plasmodium yoelii</i>
Emax_0440	7	7,81E-01	3,19E-02	0	<i>Similar to hypothetical protein PY03971</i>	XP_724205.1	<i>Plasmodium yoelii</i>
Emax_0546	7	3,61E-02	3,19E-02	0	<i>Similar to Sm protein LSm1</i>	XP_001692854.1	<i>Chlamydomonas reinhardtii</i>
Emax_0585	7	3,61E-02	3,19E-02	0	<i>Similar to hypothetical protein RUMto R_00170</i>	ZP_01966631.1	<i>Ruminococcus torques</i>
Emax_0995	7	3,61E-02	3,19E-02	0	<i>Similar to RER1 protein, putative</i>	EEA07539.1	<i>Cryptosporidium muris</i>
Emax_1031	7	3,61E-02	3,19E-02	0	<i>Similar to aldo-keto reductase</i>	XP_640007.1	<i>Dictyostelium discoideum</i>
Emax_1089	7	3,61E-02	3,19E-02	0	<i>Similar to hypothetical protein TP03_0201</i>	XP_763219.1	<i>Theileria parva</i>
Emax_0612	8	1,64E-01	3,17E-02	0	<i>Similar to calmodulin-like domain protein kinase isoform 2</i>	ABA60893.1	<i>Eimeria tenella</i>
Emax_0877	8	1,54E-01	3,17E-02	0	<i>Similar to hypothetical protein, conserved in Apicomplexan species</i>	CAQ38736.1	<i>Plasmodium knowlesi</i>
Emax_1189	8	5,43E-01	3,17E-02	0	<i>Similar to KC1_EIMTE RecName: Full=Casein kinase I</i>	AAS46021.1	<i>Eimeria tenella</i>
Emax_0841	10	4,59E-01	3,07E-02	0	<i>Similar to hypothetical protein, conserved</i>	XP_953221.1	<i>Theileria annulata</i>

Tabela 10- *Contigs* não diferencialmente expressos entre os estágios estudados que apresentaram resultado positivo de BLAST. Estão apresentados na tabela a identificação dos *contigs*, o poder de expressão (PE), ou seja, soma do número de leituras de cada estágio em cada *contig* obtido, os valores de p (*p-value*), alpha (nível crítico), *score*, resultados da busca de similaridade por BLAST contra a base de dados não redundante, os códigos de acesso das sequências e o organismo.

(Conclusão)

ID do cDNA	PE	<i>p-value</i>	alpha	<i>score</i>	Resultado da anotação automática	Código de acesso	Organismo
Emax_1108	10	7,77E-02	3,07E-02	0	<i>Similar to SPANX-N2 locus variant 1</i>	ABC61862.1	<i>Homo sapiens</i>
Emax_1153	10	3,38E-02	3,07E-02	0	<i>Similar to antigen</i>	AAA29080.1	<i>Eimeria acervulina</i>
Emax_0324	13	5,53E-02	2,85E-02	0	<i>Similar to conserved protein, COG SSU ribosomal protein S8E</i>	XP_626954.1	<i>Cryptosporidium parvum</i>
Emax_1004	13	5,17E-01	2,85E-02	0	<i>Similar to microneme protein MIC3</i>	AAR87666.1	<i>Eimeria tenella</i>
Emax_0729	14	8,52E-02	2,77E-02	0	<i>Similar to microneme protein MIC3</i>	AAR87666.1	<i>Eimeria tenella</i>
Emax_0908	14	4,02E-01	2,77E-02	0	<i>Similar to mannito l-1-phosphate dehydrogenase</i>	AAD02688.1	<i>Eimeria tenella</i>
Emax_0608	44	7,36E-02	1,19E-02	0	<i>Similar to GH15714</i>	XP_001983203.1	<i>Drosophila grimshawi</i>

Dos 260 *contigs* diferencialmente expressos classificados como proteínas hipotéticas, 28 *contigs* apresentaram motivos proteicos identificados pelo Interpro, Tabela11.

Tabela 11- Motivos proteicos identificados pelo Interpro para *contigs* diferencialmente expressos entre as diferentes fases da esporulação em *E.maxima* cujo produto foi classificado como proteína hipotética. Estão apresentados na tabela a identificação dos transcritos e os motivos proteicos.

ID do cDNA	Motivos proteicos
Emax_0060	Endonuclease/exonuclease/phosphatase
Emax_0113	Growth factor, receptor
Emax_0116	DNA polymerase III clamp loader subunit, C-terminal / Replication factor C
Emax_0127	Vesicle transport v-NSARE
Emax_0144	Apple
Emax_0170	Ribosome associated membrane RAMP4
Emax_0173	Fibronectin, type III-like fold
Emax_0195	HAD-superfamily hydrolase, subfamily IIB
Emax_0197	ARF/SAR superfamily
Emax_0270	Ankyrin
Emax_0278	Vesicle transport v-NSARE
Emax_0315	Protein kinase, core
Emax_0321	Protein kinase, core
Emax_0337	Iron hydrogenase
Emax_0398	Proteinase inhibitor I14/I15, hirudin/antistatin
Emax_0512	RNA-directed DNA polymerase (reverse transcriptase)
Emax_0539	Gurmarin-like inhibitors/Antifungal toxin
Emax_0551	C-type lectin / Aldehyde dehydrogenase, conserved site
Emax_0564	Zinc finger, C2H2-type
Emax_0570	Double-stranded RNA binding
Emax_0634	Protein kinase
Emax_0640	SFT2-like
Emax_0670	Phosphatidic acid phosphatase type 2/haloperoxidase
Emax_0685	Phosphopantetheine attachment site
Emax_0734	Protein-L-isoaspartate(D-aspartate) O-methyltransferase
Emax_0740	Sterol-sensing 5TM box
Emax_1143	Cellular retinaldehyde-binding/triple function,
Emax_1191	Insulin/IGF/relaxin

Conforme podemos observar na Tabela 11, existem três transcritos com motivos proteicos de proteínas quinases, dois transcritos relacionados ao transporte vesicular e demais transcritos de diferentes motivos proteicos.

Analisando os resultados obtidos pela busca de domínios conservados utilizando o programa RPS-BLAST, para os 260 *contigs* classificados como proteínas hipotéticas, observa-se que apenas 39 apresentaram resultados positivos (Tabela 12).

Tabela 12- Resultados positivos do programa RPS-BLAST para *contigs* diferencialmente expressos cujo produto foi classificado como proteína hipotética. Estão apresentados na tabela a identificação dos transcritos e os resultados da busca pelo programa RPS-BLAST.

(Continua)

ID do cDNA	Resultados obtidos pelo programa RPS-BLAST
Emax_0321	- <i>S_TKc, Serine/Threonine protein kinases, catalytic domain</i>
Emax_0613	<i>AAA ATPase containing von Willebrand factor type A</i>
Emax_0842	<i>aceF, dihydrolipoamide acetyltransferase</i>
Emax_0197	<i>Arf_Arl, Arf (ADP-ribosylation factor)/Arl (Arf-like) small</i>
Emax_1191	<i>BAF1_ABF1, BAF1 / ABF1 chromatin reorganising factor</i>
Emax_0870	<i>ClpA, ATPases with chaperone activity, ATP-binding subunit</i>
Emax_0925	<i>COesterase, Carboxylesterase</i>
Emax_1222	<i>CRPV_capsid, CRPV capsid protein like</i>
Emax_0875	<i>DNA translocase</i>
Emax_1143	<i>Domain in homologues of a S</i>
Emax_0116	<i>fator c de replicação</i>
Emax_0195	<i>Hydrolase_3, haloacid dehalogenase-like hydrolase</i>
Emax_0614	<i>MDN1, AAA ATPase containing von Willebrand factor type A</i>
Emax_1107	<i>MDN1, AAA ATPase containing von Willebrand factor type A</i>
Emax_0670	<i>PAP2_wunen, PAP2, wunen subfamily</i>
Emax_0740	<i>Patched, Patched family</i>
Emax_0950	<i>PCBP_like_KH, K homology RNA-binding domain</i>
Emax_0734	<i>PCMT, Protein-L-isoaspartate(D-aspartate)</i>
Emax_0326	<i>Predicted membrane protein</i>
Emax_0321	- <i>S_TKc, Serine/Threonine protein kinases, catalytic domain</i>
Emax_0399	<i>proteasome_beta_type_4, proteasome beta type-4 subunit</i>
Emax_0170	<i>RAMP4, Ribosome associated membrane protein RAMP4</i>

Tabela 12- Resultados positivos do programa RPS-BLAST para *contigs* diferencialmente expressos cujo produto foi classificado como proteína hipotética. Estão apresentados na tabela a identificação dos transcritos e os resultados da busca pelo programa RPS-BLAST.

(Conclusão)

ID do cDNA	Resultados obtidos pelo programa RPS-BLAST
Emax_0806	<i>RAP1, Rhopty-associated protein 1</i>
Emax_0354	<i>RAP1, Rhopty-associated protein 1 (RAP-1)</i>
Emax_0926	<i>REJ, REJ domain</i>
Emax_0512	<i>RNA-dependent RNA polymerase</i>
Emax_0703	<i>rpsA, 30S ribosomal protein</i>
Emax_0315	<i>S_TKc, Serine/Threonine protein kinases, catalytic domain</i>
Emax_0634	<i>Serine/Threonine protein kinases, catalytic domain</i>
Emax_0640	<i>SFT2-like protein</i>
Emax_0745	<i>Small-conductance mechanosensitive channel</i>
Emax_0306	<i>Small-conductance mechanosensitive channel</i>
Emax_0419	<i>Small-conductance mechanosensitive channel</i>
Emax_1038	<i>tolA, cell envelope integrity inner membrane3 protein TolA</i>
Emax_0422	<i>TT_ORF1, TT viral orf 1</i>
Emax_0434	<i>TT_ORF1, TT viral orf 1</i>
Emax_0887	<i>Uncharacterized protein conserved in bacteria</i>
Emax_0482	<i>V-ATPase_G, Vacuolar (H⁺)-ATPase G subunit</i>
Emax_0127	<i>V-SNARE, Vesicle transport v-SNARE protein</i>

Conforme podemos notar na Tabela 12, existem cinco transcritos similares a ATPase, três a canais mecano-sensíveis de baixa condutância, três a proteínas quinases, dois a proteínas de roptrias, dois a ORF1 de TT vírus (TTV), dois ao transporte vesicular e demais transcritos relacionados a proteínas de funções diversas.

Ao analisar conjuntamente os resultados obtidos pelas buscas no Interpro e RPS-BLAST, observa-se que apenas 16 *contigs* apresentaram resultados positivos para as duas buscas. Destes, 81,25% (13) dos transcritos foram relacionados às mesmas funções.

4.16 *Heat Map*

Os perfis de expressão gênica dos transcritos diferencialmente expressos foram visualizados utilizando a representação gráfica *Heat Map*, onde os valores de expressão são representados por cores que variam em função da intensidade da expressão. Para visualizar os diferentes níveis de expressão foi atribuída a cor azul, onde as tonalidades mais escuras representam um maior nível de transcrição do gene.

Analisando a Figura 28A observa-se que o perfil de transcrição gênica é diferenciado entre os estágios estudados. A maioria (65%) dos transcritos estava presente em um único estágio. Dentre os 379 transcritos analisados, 48 foram observados somente em Ou, 83 em Op e 116 em Os, indicando um perfil transcricional estágio específico.

Dos transcritos analisados, somente 47 (12,4%) eram compartilhados entre os três estágios estudados. Destes, 14 (3,7%) estavam presentes em Ou e Os, 25 (6,7%) em Op e Os e 46 (12,13%) em Ou e Op, mostrando que o estágio Ou parece estar mais relacionado ao estágio Op (Figura 28 B).

Analisando os resultados da anotação automática foi observado que, independentemente do estágio estudado, há uma grande quantidade de proteínas hipotéticas. Dentre as proteínas de função conhecida, a maioria está relacionada ao metabolismo.

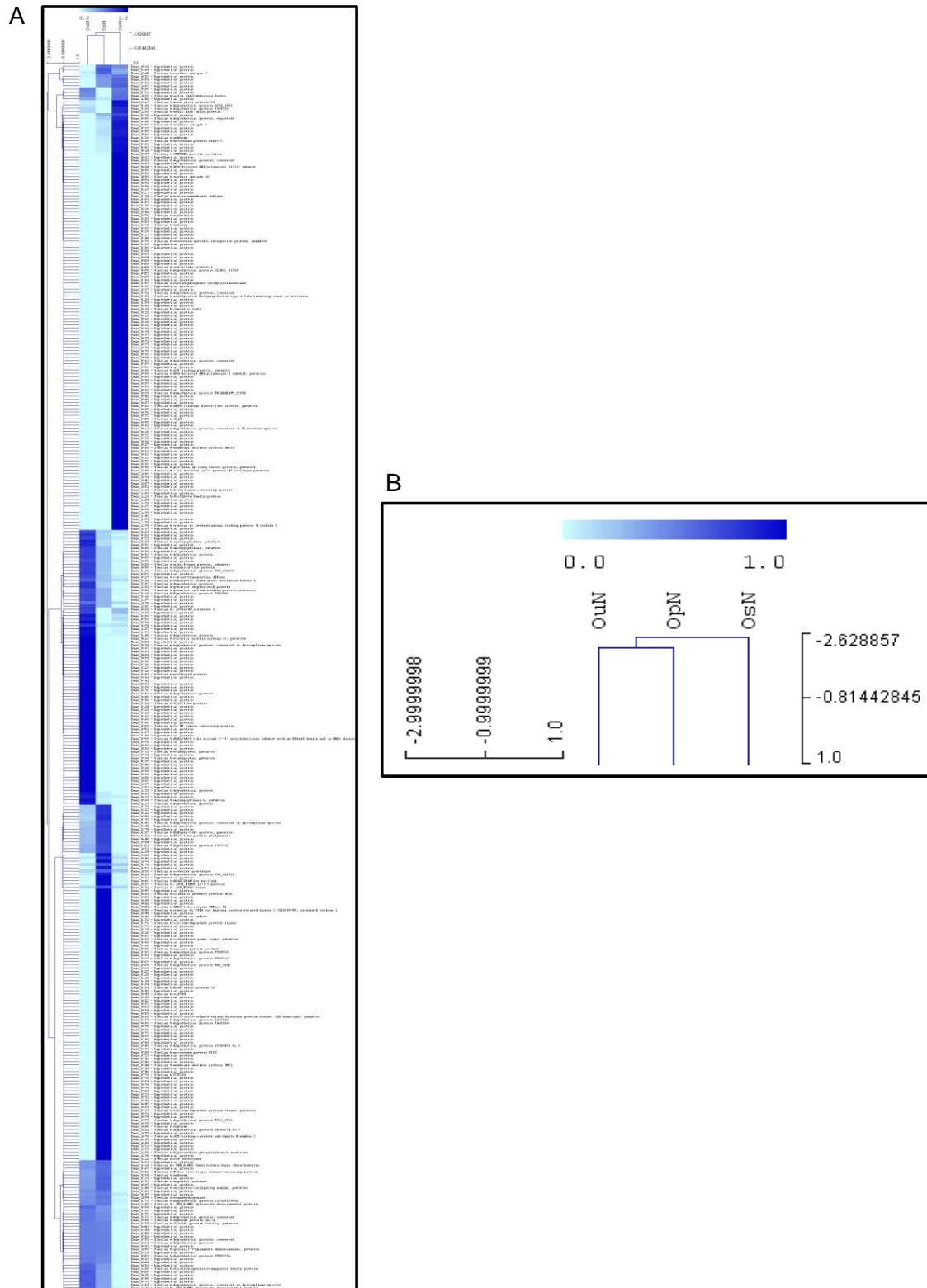


Figura 28- Comparação dos perfis de expressão gênica de três fases da esporulação de oocistos de *Eimeria maxima*. (A) *Heat map* dos níveis de expressão de 379 genes (linhas) que apresentaram níveis de transcrição diferencial entre oocistos não esporulados (Ou), parcialmente esporulados (Op) e esporulados (Os). Para cada transcrito, os valores relativos da expressão gênica são representados pelas diferentes tonalidades de azul, onde o mais escuro indica maior expressão e os tons mais claros menor expressão. (B) cabeçalho ampliado da Figura A.

4.17 Análise dos 100 transcritos diferencialmente expressos mais abundantes

Analisando os 100 transcritos diferencialmente expressos mais abundantes observa-se que a quantidade de leituras totais por *contig* (PE) variou de 23 a 600 (Tabela 13).

Grande parte dos transcritos (37%) era proveniente de somente um estágio, os quais eram representados principalmente por leituras de Os (35 transcritos); 30% eram oriundos de dois estágios e 33% dos três estágios estudados.

Os resultados de anotação automática mostraram que dentre os 100 transcritos selecionados, 68 foram classificados como proteínas hipotéticas, 19 como proteínas de função conhecida, 12 como proteínas hipotéticas conservadas e um transcrito que não teve ORF predita.

Das funções proteicas inferidas após a anotação automática, podemos observar que há três transcritos similares a antígenos de superfície, dois referentes à proteína de desenvolvimento do esporozoítos, proteína de micronema, um à proteína de choque térmico, actina, fator depolimerizante de actina e o restante das proteínas conhecidas eram relacionados ao metabolismo, dentre estas transhidrogenase, aminopeptidase e aspartil protease.

Das proteínas hipotéticas, apenas 11 apresentaram resultado positivo para a busca para motivos proteicos pelo programa RPS-BLAST e 6 para a busca de domínios conservados pelo programa InterPro. Ao analisar conjuntamente os resultados obtidos pelas buscas no Interpro e RPS-BLAST observa-se que apenas 6 *contigs* apresentaram resultados positivos para as duas buscas. Destes, 83,33% (5) dos transcritos foram relacionados às mesmas funções.

Tabela 13- Relação dos 100 *contigs* diferencialmente expressos entre as três fases da esporulação de *Eimeria maxima*, que eram mais abundantes. Estão apresentados a identificação dos *contigs*, o número de contagens obtido em oocistos não esporulados (Ou), parcialmente esporulados (Op) e esporulados (Os); o poder de expressão (PE) ou soma do número de leituras de cada estágio; resultados da busca de similaridade por BLAST contra a base de dados nr, os códigos de acesso das sequências e o organismo.

(Continua)

ID do cDNA	Ou	Op	Os	PE	Resultado da anotação automática	Código de acesso	Organismo
Emax_0398	0	29	571	600	<i>hypothetical protein</i>	-----	-----
Emax_1094	171	225	27	423	<i>Similar to transhydrogenase</i>	AAA29077.1	<i>Eimeria tenella</i>
Emax_0859	0	0	288	288	<i>hypothetical protein</i>	-----	-----
Emax_0264	90	150	25	265	<i>Similar to unknown</i>	-----	-----
Emax_0526	0	7	255	262	<i>Similar to unknown</i>	ABQ41432.1	<i>Eimeria tenella</i>
Emax_0069	0	18	237	255	<i>Similar to hypothetical protein, expressed</i>	CAJ20353.1	<i>Toxoplasma gondii</i>
Emax_1137	0	0	235	235	<i>hypothetical protein</i>	-----	-----
Emax_1190	93	111	16	220	<i>Similar to SDP_EIMBO Sporozoite developmental protein</i>	P42789	<i>Eimeria bovis</i>
Emax_1062	0	38	175	213	<i>hypothetical protein</i>	-----	-----
Emax_0527	0	0	191	191	<i>hypothetical protein</i>	-----	-----
Emax_0880	92	80	17	189	<i>Similar to unknown protein RB1-a</i>	AAC26122.1	<i>Eimeria acervulina</i>
Emax_0150	124	39	3	166	<i>Similar to zinc-finger protein, putative</i>	CAQ38960.1	<i>Plasmodium knowlesi</i>
Emax_0614	0	3	144	147	<i>hypothetical protein</i>	-----	-----
Emax_1111	73	67	7	147	<i>hypothetical protein</i>	-----	-----
Emax_0467	102	30	9	141	<i>hypothetical protein</i>	-----	-----
Emax_0247	0	1	137	138	<i>Similar to SERPIN1 protein precursor</i>	CAI72624.1	<i>Eimeria tenella</i>
Emax_1166	79	53	3	135	<i>Similar to folate/biopterin transporter family protein</i>	EEA06295.1	<i>Cryptosporidium muris</i>
Emax_0751	0	25	108	133	<i>hypothetical protein</i>	-----	-----
Emax_0962	0	0	131	131	<i>hypothetical protein</i>	-----	-----
Emax_1232	64	39	7	110	<i>Similar to SDP_EIMBO Sporozoite developmental protein</i>	P42789	<i>Eimeria bovis</i>
Emax_0029	50	47	1	98	<i>hypothetical protein</i>	-----	-----
Emax_0639	37	59	2	98	<i>Similar to aspartyl protease</i>	CAA80843.1	<i>Eimeria acervulina</i>

Tabela 13- Relação dos 100 *contigs* diferencialmente expressos entre as três fases da esporulação de *Eimeria maxima*, que eram mais abundantes. Estão apresentados a identificação dos *contigs*, o número de contagens obtido em oocistos não esporulados (Ou), parcialmente esporulados (Op) e esporulados (Os); o poder de expressão (PE) ou soma do número de leituras de cada estágio; resultados da busca de similaridade por BLAST contra a base de dados nr, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	Ou	Op	Os	PE	Resultado da anotação automática	Código de acesso	Organismo
Emax_0142	21	66	5	92	<i>Similar to hypothetical protein, conserved in Apicomplexan species</i>	CAQ42141.1	<i>Plasmodium knowlesi</i>
Emax_1038	40	22	30	92	<i>hypothetical protein</i>	-----	-----
Emax_0388	42	34	12	88	<i>hypothetical protein</i>	-----	-----
Emax_0915	0	27	60	87	<i>hypothetical protein</i>	-----	-----
Emax_0361	64	14	6	84	<i>hypothetical protein</i>	-----	-----
Emax_0631	0	0	83	83	<i>hypothetical protein</i>	-----	-----
Emax_0130	0	0	79	79	<i>Similar to non-transmembrane antigen</i>	AAO65977.1	<i>Toxoplasma gondii</i>
Emax_0152	56	17	6	79	<i>Similar to hypothetical protein PVX_090030</i>	XP_001615098.1	<i>Plasmodium vivax</i>
Emax_0435	0	76	0	76	<i>Similar to hypothetical protein MGL_0149</i>	XP_001732374.1	<i>Malassezia globosa</i>
Emax_0572	37	31	7	75	<i>hypothetical protein</i>	-----	-----
Emax_0812	37	33	2	72	<i>hypothetical protein</i>	-----	-----
Emax_0963	0	0	72	72	<i>hypothetical protein</i>	-----	-----
Emax_0394	0	13	57	70	<i>hypothetical protein</i>	-----	-----
Emax_0297	51	13	2	66	<i>Similar to hypothetical protein</i>	XP_980864.1	<i>Mus musculus</i>
Emax_0692	1	14	48	63	<i>hypothetical protein</i>	-----	-----
Emax_0961	0	0	63	63	<i>hypothetical protein</i>	-----	-----
Emax_1095	30	0	32	62	<i>hypothetical protein</i>	-----	-----
Emax_0517	38	13	9	60	<i>Similar to cation-transporting ATPase</i>	ACB97673.1	<i>Eimeria acervulina</i>
Emax_0752	49	8	3	60	<i>hypothetical protein</i>	-----	-----
Emax_0983	0	0	60	60	<i>hypothetical protein</i>	-----	-----
Emax_0337	0	0	56	56	<i>hypothetical protein</i>	-----	-----

Tabela 13- Relação dos 100 *contigs* diferencialmente expressos entre as três fases da esporulação de *Eimeria maxima*, que eram mais abundantes. Estão apresentados a identificação dos *contigs*, o número de contagens obtido em oocistos não esporulados (Ou), parcialmente esporulados (Op) e esporulados (Os); o poder de expressão (PE) ou soma do número de leituras de cada estágio; resultados da busca de similaridade por BLAST contra a base de dados nr, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	Ou	Op	Os	PE	Resultado da anotação automática	Código de acesso	Organismo
Emax_0933	0	0	56	56	<i>hypothetical protein</i>	-----	-----
Emax_0197	13	0	39	52	<i>hypothetical protein</i>	-----	-----
Emax_0338	12	0	40	52	<i>hypothetical protein</i>	-----	-----
Emax_1165	32	19	1	52	<i>Similar to hypothetical protein, conserved in Apicomplexan species</i>	CAQ38339.1	<i>Plasmodium knowlesi</i>
Emax_0113	51	0	0	51	<i>hypothetical protein</i>	-----	-----
Emax_0918	0	0	51	51	<i>hypothetical protein</i>	-----	-----
Emax_0521	18	31	0	49	<i>hypothetical protein</i>	-----	-----
Emax_0052	0	0	48	48	<i>hypothetical protein</i>	-----	-----
Emax_0285	0	0	48	48	<i>hypothetical protein</i>	-----	-----
Emax_1143	8	2	38	48	<i>hypothetical protein</i>	-----	-----
Emax_0302	0	1	46	47	<i>hypothetical protein</i>	-----	-----
Emax_0331	5	24	15	44	<i>Similar to ACT_to XGO Actin</i>	AAC13766.1	<i>Toxoplasma gondii</i>
Emax_0558	0	1	43	44	<i>hypothetical protein</i>	-----	-----
Emax_0845	0	0	44	44	<i>hypothetical protein</i>	-----	-----
Emax_0109	38	1	4	43	<i>Similar to hypothetical protein</i>	XP_001238682.1	<i>Eimeria tenella</i>
Emax_0955	40	2	1	43	<i>Similar to aminopeptidase n, putative</i>	XP_954343.1	<i>Theileria annulata</i>
Emax_0956	0	2	41	43	<i>hypothetical protein</i>	-----	-----
Emax_0208	0	1	41	42	<i>Similar to microneme protein Etmic-2</i>	AAD05565.1	<i>Eimeria tenella</i>
Emax_0640	3	26	12	41	<i>hypothetical protein</i>	-----	-----
Emax_0017	1	0	38	39	<i>Similar to heat shock protein 90</i>	AAS17969.1	<i>Eimeria acervulina</i>
Emax_0483	19	18	2	39	<i>Similar to hypothetical protein PFE0220w</i>	XP_001351603.1	<i>Plasmodium falciparum</i>

Tabela 13- Relação dos 100 *contigs* diferencialmente expressos entre as três fases da esporulação de *Eimeria maxima*, que eram mais abundantes. Estão apresentados a identificação dos *contigs*, o número de contagens obtido em oocistos não esporulados (Ou), parcialmente esporulados (Op) e esporulados (Os); o poder de expressão (PE) ou soma do número de leituras de cada estágio; resultados da busca de similaridade por BLAST contra a base de dados nr, os códigos de acesso das sequências e o organismo.

(Continuação)

ID do cDNA	Ou	Op	Os	PE	Resultado da anotação automática	Código de acesso	Organismo
Emax_0501	0	0	38	38	<i>hypothetical protein</i>	-----	-----
Emax_0153	7	28	0	35	<i>hypothetical protein</i>	-----	-----
Emax_0399	0	0	35	35	<i>hypothetical protein</i>	-----	-----
Emax_1088	4	16	15	35	<i>Similar to cysteine proteinase</i>	AAL60053.1	<i>Toxoplasma gondii</i>
Emax_0632	0	0	33	33	<i>hypothetical protein</i>	-----	-----
Emax_0281	0	30	2	32	<i>hypothetical protein</i>	-----	-----
Emax_0434	0	0	32	32	<i>hypothetical protein</i>	-----	-----
Emax_0530	1	0	30	31	<i>Similar to hypothetical protein PF14_0571</i>	XP_001348745.1	<i>Plasmodium falciparum</i>
Emax_0615	0	0	31	31	<i>hypothetical protein</i>	-----	-----
Emax_0872	0	0	31	31	<i>hypothetical protein</i>	-----	-----
Emax_0090	0	0	30	30	<i>Similar to surface antigen 10</i>	ABR53732.1	<i>Eimeria tenella</i>
Emax_0148	13	17	0	30	<i>hypothetical protein</i>	-----	-----
Emax_0287	0	1	29	30	<i>Similar to surface antigen 9</i>	CAE52300.1	<i>Eimeria tenella</i>
Emax_0982	0	0	30	30	<i>hypothetical protein</i>	-----	-----
Emax_0170	0	0	29	29	<i>hypothetical protein</i>	-----	-----
Emax_0482	0	0	29	29	<i>hypothetical protein</i>	-----	-----
Emax_0325	0	0	28	28	<i>Similar to unknown</i>	ABQ41434.1	<i>Eimeria tenella</i>
Emax_0728	0	1	27	28	<i>hypothetical protein</i>	-----	-----
Emax_1020	6	0	22	28	<i>Similar to actin depolymerizing facto r</i>	ABM89551.1	<i>Eimeria tenella</i>

Tabela 13- Relação dos 100 *contigs* diferencialmente expressos entre as três fases da esporulação de *Eimeria maxima*, que eram mais abundantes. Estão apresentados a identificação dos *contigs*, o número de contagens obtido em oocistos não esporulados (Ou), parcialmente esporulados (Op) e esporulados (Os); o poder de expressão (PE) ou soma do número de leituras de cada estágio; resultados da busca de similaridade por BLAST contra a base de dados nr, os códigos de acesso das sequências e o organismo.
(Conclusão)

ID do cDNA	Ou	Op	Os	PE	Resultado da anotação automática	Código de acesso	Organismo
Emax_0638	0	0	27	27	<i>hypothetical protein</i>	-----	-----
Emax_0903	0	0	27	27	<i>hypothetical protein</i>	-----	-----
Emax_0013	0	0	26	26	<i>hypothetical protein</i>	-----	-----
Emax_0106	0	1	25	26	<i>hypothetical protein</i>	-----	-----
Emax_0278	3	12	11	26	<i>hypothetical protein</i>	-----	-----
Emax_0430	0	0	26	26	-----	-----	-----
Emax_0889	24	2	0	26	<i>hypothetical protein</i>	-----	-----
Emax_0219	0	3	22	25	<i>hypothetical protein</i>	-----	-----
Emax_0321	16	4	5	25	<i>hypothetical protein</i>	-----	-----
Emax_1012	1	10	14	25	<i>Similar to surface antigen 4</i>	CAE52296.1	<i>Eimeria tenella</i>
Emax_0363	10	14	0	24	<i>hypothetical protein</i>	-----	-----
Emax_0613	0	0	23	23	<i>hypothetical protein</i>	-----	-----
Emax_0847	0	22	1	23	<i>hypothetical protein</i>	-----	-----

4.17.1 Comparação dos produtos anotados de *E. maxima* com sequências de outros parasitas do Filo Apicomplexa.

Dos 100 transcritos selecionados, 52 apresentaram similaridade com sequências depositadas nos bancos de dados ToxoDb, CryptoDb e PlasmoDB.

Após a busca por similaridade em cada banco observou-se que 55 transcritos eram similares a *Toxoplasma gondii*, 33 a *Plasmodium* spp., 27 a *Cryptosporidium* spp. e 5 a *Neospora caninum*. Analisando comparativamente os produtos de função conhecida, observou-se que os resultados obtidos eram condizentes com os da anotação automática. Das proteínas classificadas como hipotéticas pela anotação automática, foi observado pela presente busca que 16 eram similares à proteínas hipotéticas conservadas entre as espécies, 2 à proteínas quinases, e demais à miosina, proteína de parede de oocisto, proteína contendo o domínio BTB/PO2, proteína de sinalização celular CRAL/TRIO, proteína contendo domínio KH, subunidade ribossomal do RNA, proteína L-isoaspartato O-metiltransferase e fator ADP-ribosilação. Para estes transcritos, esta busca possibilitou um incremento dos resultados da anotação.

4.17.1.1 Comparação dos produtos anotados de *E. maxima* com sequências ORESTES de *E. acervulina* e ORESTES/ESTs de *E. tenella*

Ao comparar as 100 sequências reconstruídas diferencialmente expressas mais abundantes de *Eimeria maxima* com as sequências ORESTES e/ou ESTs de *E. acervulina* e *E. tenella* (Tabela 14) obtivemos respectivamente 56 e 59 sequências similares, dos quais 45 transcritos eram semelhantes entre as três espécies. Dentre os transcritos similares, foi observado um transcrito similar entre as três espécies comparadas, 9 transcritos de *E. acervulina* e 12 de *E. tenella* com perfil de expressão semelhante à *E. maxima* para os estágios da esporulação (Tabela 15).

Tabela 14- *Contigs* diferencialmente expressos entre os 100 transcritos mais abundantes de *E. maxima* cujos produtos foram similares aos transcritos de *E. tenella* e *E. acervulina*. Estão apresentados na tabela a identificação dos transcritos de *E. maxima*, os resultados da busca de similaridade por BLAST contra a base de dados nr, a identificação dos transcritos de *E. tenella* e de *E. acervulina*.

(Continua)

ID de <i>E. maxima</i>	Resultado da anotação automática	ID de <i>E. tenella</i>	ID de <i>E. acervulina</i>
Emax_0017	<i>Similar to heat shock protein 90</i>	Eten_0346	Eace_1254
Emax_0052	<i>hypothetical protein</i>	Eten_6114	*ND
Emax_0069	<i>Similar to hypothetical protein, expressed</i>	Eten_0827	Eace_0243
Emax_0109	<i>Similar to hypothetical protein</i>	Eten_1222	*ND
Emax_0113	<i>hypothetical protein</i>	Eten_0561	Eace_0843
Emax_0130	<i>Similar to non-transmembrane antigen</i>	*ND	Eace_0315
Emax_0142	<i>Similar to hypothetical protein, conserved in Apicomplexan species</i>	Eten_0500	Eace_1264
Emax_0148	<i>hypothetical protein</i>	Eten_0562	Eace_0409
Emax_0150	<i>Similar to zinc-finger protein, putative</i>	Eten_0620	Eace_1275
Emax_0152	<i>Similar to hypothetical protein PVX_090030</i>	Eten_0599	*ND
Emax_0197	<i>hypothetical protein</i>	Eten_2865	Eace_0895
Emax_0208	<i>Similar to microneme protein Etmic-2</i>	Eten_0776	*ND
Emax_0247	<i>Similar to SERPIN1 protein precursor</i>	Eten_3216	Eace_0134
Emax_0264	<i>Similar to unknown</i>	Eten_1237	Eace_0293
Emax_0278	<i>hypothetical protein</i>	Eten_0811	Eace_0051
Emax_0281	<i>hypothetical protein</i>	Eten_7546	Eace_0879
Emax_0285	<i>hypothetical protein</i>	Eten_1588	Eace_0133
Emax_0297	<i>Similar to hypothetical protein</i>	Eten_1213	Eace_0414

*ND- não foi encontrado transcrito significativamente similar ao de *Eimeria maxima*

Tabela 14- *Contigs* diferencialmente expressos entre os 100 transcritos mais abundantes de *E. maxima* cujos produtos foram similares aos transcritos de *E. tenella* e *E. acervulina*. Estão apresentados na tabela a identificação dos transcritos de *E. maxima*, os resultados da busca de similaridade por BLAST contra a base de dados nr, a identificação dos transcritos de *E. tenella* e de *E. acervulina*.

(Continuação)

ID de <i>E. maxima</i>	Resultado da anotação automática	ID de <i>E. tenella</i>	ID de <i>E. acervulina</i>
Emax_0302	<i>hypothetical protein</i>	Eten_2449	Eace_2865
Emax_0321	<i>hypothetical protein</i>	Eten_2741	*ND
Emax_0331	<i>Similar to ACT_to XGO Actin</i>	Eten_2546	Eace_0998
Emax_0337	<i>hypothetical protein</i>	*ND	Eace_0934
Emax_0338	<i>hypothetical protein</i>	Eten_0827	Eace_0243
Emax_0361	<i>hypothetical protein</i>	Eten_2308	*ND
Emax_0388	<i>hypothetical protein</i>	Eten_1526	Eace_0735
Emax_0394	<i>hypothetical protein</i>	Eten_2750	Eace_1165
Emax_0398	<i>hypothetical protein</i>	Eten_3381	Eace_0144
Emax_0467	<i>hypothetical protein</i>	*ND	Eace_0526
Emax_0483	<i>Similar to hypothetical protein PFE0220w</i>	Eten_0827	Eace_0243
Emax_0501	<i>hypothetical protein</i>	Eten_6186	Eace_0425
Emax_0515	<i>Similar to hypothetical protein PY03482</i>	Eten_1575	*ND
Emax_0517	<i>Similar to cation-transporting ATPase</i>	Eten_1565	*ND
Emax_0526	<i>Similar to unknown</i>	*ND	Eace_0264
Emax_0527	<i>hypothetical protein</i>	*ND	Eace_1051
Emax_0558	<i>hypothetical protein</i>	Eten_3005	Eace_3133
Emax_0572	<i>hypothetical protein</i>	Eten_1473	Eace_0224

*ND- não foi encontrado transcrito significativamente similar ao de *Eimeria maxima*

Tabela 14- *Contigs* diferencialmente expressos entre os 100 transcritos mais abundantes de *E. maxima* cujos produtos foram similares aos transcritos de *E. tenella* e *E. acervulina*. Estão apresentados na tabela a identificação dos transcritos de *E. maxima*, os resultados da busca de similaridade por BLAST contra a base de dados nr, a identificação dos transcritos de *E. tenella* e de *E. acervulina*.
(Continuação)

ID de <i>E. maxima</i>	Resultado da anotação automática	ID de <i>E. tenella</i>	ID de <i>E. acervulina</i>
Emax_0614	<i>hypothetical protein</i>	Eten_0063	Eace_0360
Emax_0631	<i>hypothetical protein</i>	Eten_3127	*ND
Emax_0639	<i>Similar to aspartyl protease</i>	Eten_2297	Eace_0021
Emax_0640	<i>hypothetical protein</i>	Eten_2547	Eace_1476
Emax_0692	<i>hypothetical protein</i>	Eten_0057	Eace_1003
Emax_0751	<i>hypothetical protein</i>	Eten_0198	Eace_0879
Emax_0812	<i>hypothetical protein</i>	Eten_1166	*ND
Emax_0828	<i>hypothetical protein</i>	Eten_6653	Eace_0385
Emax_0845	<i>hypothetical protein</i>	Eten_2289	Eace_1005
Emax_0847	<i>hypothetical protein</i>	*ND	Eace_0376
Emax_0859	<i>hypothetical protein</i>	*ND	Eace_0398
Emax_0880	<i>Similar to unknown protein RB1-a</i>	Eten_1975	Eace_1098
Emax_0889	<i>hypothetical protein</i>	*ND	Eace_0753
Emax_0915	<i>hypothetical protein</i>	Eten_2750	Eace_1165
Emax_0918	<i>hypothetical protein</i>	Eten_3474	Eace_2235
Emax_0955	<i>Similar to aminopeptidase n, putative</i>	Eten_6195	*ND
Emax_0961	<i>hypothetical protein</i>	Eten_1586	*ND
Emax_0962	<i>hypothetical protein</i>	Eten_2592	Eace_1245

*ND- não foi encontrado transcrito significativamente similar ao de *Eimeria maxima*

Tabela 14- *Contigs* diferencialmente expressos entre os 100 transcritos mais abundantes de *E. maxima* cujos produtos foram similares aos transcritos de *E. tenella* e *E. acervulina*. Estão apresentados na tabela a identificação dos transcritos de *E. maxima*, os resultados da busca de similaridade por BLAST contra a base de dados nr, a identificação dos transcritos de *E. tenella* e de *E. acervulina*.

(Conclusão)

ID de <i>E. maxima</i>	Resultado da anotação automática	ID de <i>E. tenella</i>	ID de <i>E. acervulina</i>
Emax_0963	<i>hypothetical protein</i>	Eten_2949	*ND
Emax_0983	<i>hypothetical protein</i>	Eten_1795	*ND
Emax_1012	<i>Similar to surface antigen 4</i>	Eten_2537	Eace_0230
Emax_1020	<i>Similar to actin depolymerizing facto r</i>	Eten_3526	Eace_0936
Emax_1023	<i>hypothetical protein</i>	Eten_6343	Eace-0349
Emax_1038	<i>hypothetical protein</i>	Eten_6190	Eace_1108
Emax_1062	<i>hypothetical protein</i>	Eten_2750	Eace_1236
Emax_1088	<i>Similar to cysteine proteinase</i>	*ND	Eace_3054
Emax_1094	<i>Similar to transhydrogenase</i>	Eten_1199	Eace_0433
Emax_1095	<i>hypothetical protein</i>	Eten_3006	Eace_0990
Emax_1111	<i>hypothetical protein</i>	*ND	Eace_0863
Emax_1137	<i>hypothetical protein</i>	Eten_7075	Eace_0996
Emax_1143	<i>hypothetical protein</i>	*ND	Eace_1212
Emax_1165	<i>Similar to hypothetical protein, conserved in Apicomplexan species</i>	Eten_1165	Eace_0228
Emax_1190	<i>Similar to SDP_EIMBO Sporozoite developmental protein</i>	Eten_0516	Eace_1266
Emax_1232	<i>Similar to SDP_EIMBO Sporozoite developmental protein</i>	Eten_1291	Eace_1266

*ND- não foi encontrado transcrito significativamente similar ao de *Eimeria maxima*

Tabela 15- Relação dos *contigs* diferencialmente expressos entre os 100 mais abundantes de *E. maxima* cujos perfis de expressão gênica foram similares para pelo menos uma espécie, *E. tenella* ou *E. acervulina*. Estão apresentados na tabela a identificação dos transcritos de *E. maxima*, *E. tenella* e de *E. acervulina* e os respectivos perfis de expressão. Os estagios evolutivos avaliados foram (Ou) oocistos não esporulados, (Op) oocistos parcialmente esporulados e (Os) oocistos esporulados Os perfis de expressão semelhantes à *E. maxima* estão grafados em azul e os diferentes em laranja.

<i>E. maxima</i>		<i>E. tenella</i>		<i>E. acervulina</i>	
ID do cDNA	Perfil de expressão	ID do cDNA	Perfil de expressão	ID do cDNA	Perfil de expressão
Emax_0069	Os>Op>Ou	Eten_0827	Os>Op>Ou	Eace_0243	Op>Ou>Os
Emax_0113	Ou	Eten_0561	Ou	Eace_0843	Op>Ou>Os
Emax_0142	Op>Ou>Os	Eten_0500	Op>Ou>Os	Eace_1264	Ou>Op>Os
Emax_0150	Ou>Op>Os	Eten_0620	Ou>Op>Os	Eace_1275	Ou
Emax_0152	Ou>Op>Os	Eten_0599	Ou>Op>Os	*ND	*ND
Emax_0208	Os>Op	Eten_0776	Os>Op	*ND	*ND
Emax_0247	Os>Op	Eten_3216	Os>Ou -	Eace_0134	Os>Op
Emax_0264	Op>Ou>Os	Eten_1237	Op>Ou>Os	Eace_0293	Ou>Op>Os
Emax_0285	Os	Eten_1588	Os>Ou	Eace_0133	Os
Emax_0388	Ou>Op>Os	Eten_1526	Ou>Op>Os	Eace_0735	Ou>Os>Op
Emax_0572	Ou>Op>Os	Eten_1473	Op>Ou	Eace_0224	Ou>Op>Os
Emax_0692	Op>Os>Ou	Eten_0057	**	Eace_1003	Op>Os>Ou
Emax_0845	Os	Eten_2289	Os	Eace_1005	Os>Op
Emax_0880	Ou>Op>Os	Eten_1975	Ou	Eace_1098	Ou>Op>Os
Emax_0915	Op>Os	Eten_2750	Os	Eace_1165	Op>Os
Emax_0962	Os	Eten_2592	**	Eace_1245	Os
Emax_0983	Os	Eten_1795	Os	*ND	*ND
Emax_1062	Os>Op	Eten_2750	Os>Ou=Op-	Eace_1236	Os>Op
Emax_1094	Op>Ou>Os	Eten_1199	Op>Ou>Os	Eace_0433	Ou>Op>Os
Emax_1190	Op>Ou>Os	Eten_0516	Op>Ou>Os	Eace_1266	Ou>Op>Os
Emax_1166	Ou>Op>Os	Eten_1535	Ou>Op>Os	Eace_0634	Ou>Op>Os
Emax_1232	Ou>Op>Os	Eten_1291	Op>Ou>Os	Eace_1266	Ou>Op>Os

*ND- não foi encontrada similaridade significativa ao transcrito de *Eimeria maxima*

** não apresentou contagem nos estágios Ou, Op ou Os.

4.17.2 Análise no programa Threader 3 (*Protein Fold Recognition by Threading*)

Com a finalidade de inferir possíveis funções, as proteínas hipotéticas (resultado negativo após busca por similaridade no programa BLAST) presentes no grupo dos 100 transcritos reconstruídos diferencialmente expressos mais abundantes foram submetidas à análise no programa Threader 3.

Das 68 proteínas hipotéticas selecionadas, 49 apresentaram enovelamento predito similar com outras proteínas depositadas no banco de dados RCSB PDB (*The Protein Data Bank*, <http://www.pdb.org/pdb/home>), Tabela 16.

Ao analisar os índices de confiabilidade da predição proteica (*z-score*) foi observado que 12 proteínas apresentaram índice de confiabilidade altamente significativo (*z-score* >4,0), indicando que a predição tem alta probabilidade de estar correta; 18 proteínas apresentaram índice de confiabilidade significativo (*z-score* >3,5), indicando boas chances da predição estar correta e 19 proteínas apresentaram índice de confiabilidade médio (incerto), entre 2,7 e 3,5, sugerindo que a predição possivelmente está correta.

Analisando a tabela 16 observa-se que as proteínas identificadas apresentam várias funções, das quais a maioria é relacionada ao metabolismo. Dentre as demais proteínas, seis estão relacionadas ao proteassomo e quatro são proteínas hipotéticas de estrutura conhecida.

Tabela 16- Proteínas hipotéticas presentes nos 100 *contigs* diferencialmente expressos mais abundantes em *E. maxima* submetidas a análise no programa Threader. Estão apresentados na tabela os números de identificação dos transcritos de *E.maxima*, o valor de *z-score*, classificação do índice de confiabilidade da predição proteica (conforme o manual), número de identificação no banco de dados RCSB PDB (*The Protein Data Bank*) e a proteína identificada.

(Continua)

ID do cDNA	<i>z-score</i>	Classificação (índice de confiabilidade)	ID do PDB	Identificação da proteína
Emax_0113	4.32	Muito significativa	2ntkA0	<i>Imp cyclohydrolase</i>
Emax_0918	4.26	Muito significativa	3fpiA0	<i>2-c-methyl-d-erythritol 2,4-cyclodiphosphate</i>
Emax_0278	4.23	Muito significativa	1shuX0	<i>Anthrax toxin receptor 2</i>
Emax_1111	4.19	Muito significativa	1mldA0	<i>Malate dehydrogenase</i>
Emax_0501	4.18	Muito significativa	1iruI0	<i>20s proteasome</i>
Emax_0388	4.17	Muito significativa	1jdhA0	<i>Beta-catenin</i>
Emax_0845	4.15	Muito significativa	3kraA0	<i>Geranyl diphosphate synthase large subunit</i>
Emax_0859	4.15	Muito significativa	2a8iA0	<i>Threonine aspartase 1</i>
Emax_0197	4.14	Muito significativa	1s12A0	<i>Hypothetical protein tm1457</i>
Emax_0361	4.13	Muito significativa	1h5yA0	<i>Hisf</i>
Emax_0962	4.03	Muito significativa	3edmA0	<i>Short chain dehydrogenase</i>
Emax_1143	4.03	Muito significativa	1w0mA0	<i>Triosephosphate isomerase</i>
Emax_0640	3.97	Significante	1wx1A0	<i>Nicotinate-nucleotide--dimethylbenzimidazole</i>
Emax_1023	3.88	Significante	1qdeA0	<i>Translation initiation factor 4a</i>
Emax_0812	3.83	Significante	1iruL0	<i>20s proteasome</i>
Emax_0281	3.83	Significante	3mh5A0	<i>Protease do</i>
Emax_0467	3.74	Significante	1xqrA0	<i>HSPBP1 protein</i>
Emax_0828	3.69	Significante	2z5yA0	<i>Amine oxidase [flavin-containing] a</i>

Tabela 16- Proteínas hipotéticas presentes nos 100 *contigs* diferencialmente expressos mais abundantes em *E. maxima* submetidas a análise no programa Threader. Estão apresentados na tabela os números de identificação dos transcritos de *E.maxima*, o valor de *z-score*, classificação do índice de confiabilidade da predição proteica (conforme o manual), número de identificação no banco de dados RCSB PDB (*The Protein Data Bank*) e a proteína identificada.

(Continuação)

ID do cDNA	<i>z-score</i>	Classificação (índice de confiabilidade)	ID do PDB	Identificação da proteína
Emax_0983	3.67	Significante	1s1pA0	<i>Aldo-keto reductase family 1 member c3</i>
Emax_0631	3.67	Significante	2qk4A0	<i>Trifunctional purine biosynthetic protein</i>
Emax_0752	3.65	Significante	1jcfA0	<i>Rod shape-determining protein mreB</i>
Emax_0963	3.63	Significante	1i1qA0	<i>Anthranilate synthase component i</i>
Emax_0933	3.62	Significante	1rgzA0	<i>Class c beta-lactamase</i>
Emax_0321	3.62	Significante	3cfsB0	<i>Histone-binding protein RBBP7</i>
Emax_0052	3.61	Significante	2a87A0	<i>Thioredoxin reductase</i>
Emax_0889	3.58	Significante	3lv0A0	<i>Extragenic suppressor protein suhb</i>
Emax_0751	3.55	Significante	2jdqA0	<i>Importin alpha-1 subunit</i>
Emax_0961	3.54	Significante	1g3qA0	<i>Cell division inhibitor</i>
Emax_0148	3.52	Significante	3lkdA0	<i>Type i restriction-modification system</i>
Emax_0302	3.51	Significante	1ll7A0	<i>Chitinase 1</i>
Emax_0558	3.50	Incerto	1iruJ0	<i>20s proteasome</i>
Emax_0338	3.47	Incerto	1hqkA0	<i>6,7-dimethyl-8-ribityllumazine synthase</i>
Emax_0692	3.44	Incerto	3mka20	<i>Proteasome subunit alpha</i>
Emax_1095	3.44	Incerto	1lssA0	<i>TRK system potassium uptake protein trka homolog</i>
Emax_1137	3.44	Incerto	3onpA0	<i>tRNA/rRNA methyltransferase (spou)</i>
Emax_0527	3.40	Incerto	3kljA0	<i>NAD(FAD)-dependent dehydrogenase, nirb-family</i>
Emax_0572	3.40	Incerto	2xdwA0	<i>Dwa0 prolyl endopeptidase</i>

Tabela 16- Proteínas hipotéticas presentes nos 100 *contigs* diferencialmente expressos mais abundantes em *E. maxima* submetidas a análise no programa Threader. Estão apresentados na tabela os números de identificação dos transcritos de *E.maxima*, o valor de *z-score*, classificação do índice de confiabilidade da predição proteica (conforme o manual), número de identificação no banco de dados RCSB PDB (*The Protein Data Bank*) e a proteína identificada.

(Conclusão)

ID do cDNA	<i>z-score</i>	Classificação (índice de confiabilidade)	ID do PDB	Identificação da proteína
Emax_0029	3.37	Incerto	2z1nA0	<i>Dehydrogenase</i>
Emax_0153	3.37	Incerto	3md1A0	<i>Nuclear and cytoplasmic polyadenylated RNA-binding protein</i>
Emax_0915	3.37	Incerto	2xesA0	<i>Protein pat1 homolog 1</i>
Emax_0337	3.34	Incerto	1ekqA0	<i>Hydroxyethylthiazole kinase</i>
Emax_0678	3.26	Incerto	2ekmA0	<i>Hypothetical protein st1511</i>
Emax_0398	3.26	Incerto	2cveA0	<i>Hypothetical protein ttha1053</i>
Emax_0956	3.22	Incerto	3io3A0	<i>Deha2d07832p</i>
Emax_0363	3.17	Incerto	1yarA0	<i>Proteasome alpha subunit</i>
Emax_0847	3.03	Incerto	1oi0A0	<i>Hypothetical protein af2198</i>
Emax_1038	3.02	Incerto	2ffuA0	<i>Polypeptide n-acetylgalactosaminyltransferase 2</i>
Emax_0614	2.99	Incerto	2c3fA0	<i>Hyaluronidase, phage associated</i>
Emax_0285	2.94	Incerto	1yarH0	<i>Proteasome alpha subunit</i>
Emax_1062	2.92	Incerto	2gekA0	<i>Phosphatidylinositol mannosyltransferase (pima)</i>
Emax_0394	2.26	Incerto	2z5kA0	<i>Transportin-1</i>

4.18 PCR em tempo real (qPCR)

Os experimentos de PCR em tempo real (qPCR) foram realizados com o objetivo de validar os perfis de transcrição obtidos por *Northern* digital. Para isso, foi utilizada a metodologia de quantificação da expressão relativa que se baseia no cálculo da relação da expressão do gene de interesse (amostras experimentais) em relação à expressão do gene controle (Peirson et al., 2003).

4.18.1 Seleção de genes constitutivamente expressos (controle)

Inicialmente, foi realizada uma busca de genes controle, ou seja, que são expressos de forma constitutiva e com transcrição semelhante nas diferentes fases de esporulação de oocistos. A seleção destes genes foi feita a partir de transcritos reconstruídos, conforme critérios estabelecidos no item 3.19, e os genes selecionados estão descritos na tabela 2.

Quando comparamos os resultados obtidos da quantificação da expressão gênica por qPCR entre os candidatos à gene controle, observamos que apenas o transcrito Emax_0608, gene relacionado a ubiquitina, apresentou uma menor variação de Ct (<0,5) entre os estágios (oocistos esporulados, oocistos parcialmente esporulados e oocistos esporulados) estudados nos diferentes experimentos (Figura 29), sendo este o transcrito selecionado como gene controle.

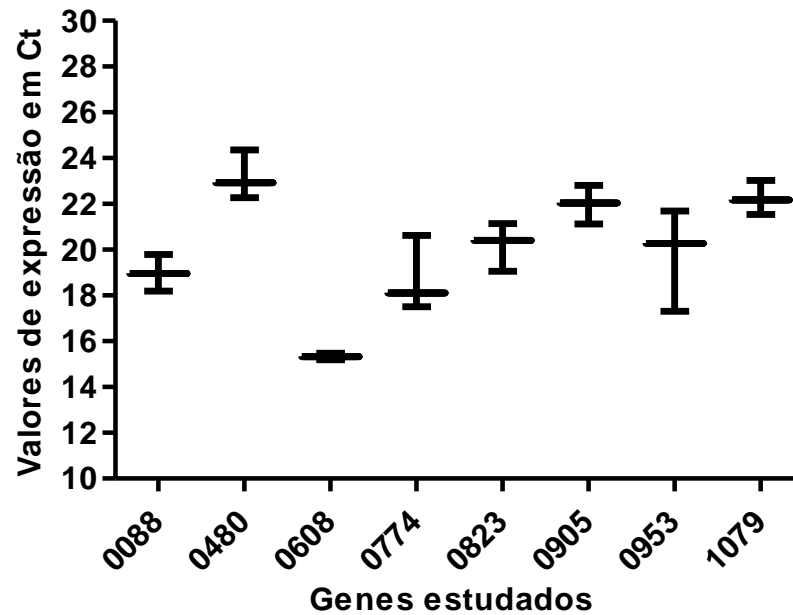


Figura 29- Valores de Ct (PCR *Cycle threshold*) para os transcritos candidatos a genes controle. Os traços centrais equivalem a média dos valores obtidos em diferentes experimentos utilizando amostras de cDNA sintetizadas a partir de RNAm de oocistos não esporulados, oocistos parcialmente esporulados e oocistos esporulados. Os traços externos correspondem aos valores mínimos e máximos obtidos nos experimentos.

4.18.2 Seleção de genes diferencialmente expressos

A seleção dos genes diferencialmente expressos estudados foi realizada a partir de transcritos reconstruídos obtidos por ORESTES segundo os critérios descritos no item 3.19 e estão listados na Tabela 17 com resultados de anotação automática. Foram selecionados genes que estavam dentro do grupo dos 100 transcritos diferencialmente expressos mais abundantes de função conhecida e genes diferencialmente expressos, cujo produto era similar à proteínas de importante função em parasitas do Filo Apicomplexa, como proteína rombóide e dinamina. Exceção feita à proteína hipotética, que por apresentar em ORESTES grande número de leituras somente para oocistos esporulados, poderia ser um marcador desta fase da esporulação.

Tabela 17- Relação dos *contigs* diferencialmente expressos entre as três fases da esporulação de *Eimeria maxima*, que foram avaliados por RT-qPCR. Estão apresentados na tabela a identificação dos *contigs*, o número de contagens obtido em oocistos não esporulados (Ou), parcialmente esporulados (Op) e esporulados (Os); valores de p; resultados da busca de similaridade por BLAST contra a base de dados nr, os domínios conservados identificados por RPS-BLAST e os motivos proteicos identificados pelo Interpro.

ID do cDNA	Perfil de expressão			<i>p-value</i>	Resultado da anotação automática (BLAST)	Domínios conservados (RSP-BLAST)	Motivos proteicos (Interpro)
	Ou	Op	Os				
Emax_0017	1	0	38	0	Proteína de choque térmico HSP90	<i>Hsp90 protein</i>	<i>Heat shock protein Hsp90</i>
Emax_0130	0	0	79	0	Antígeno não transmembranar	<i>PARG_cat, Poly (ADP-ribose) glycohydrolase (PARG)</i>	<i>Sugar transporter</i>
Emax_0208	0	1	41	0	Proteína de micronema 2	<i>Etmic-2, Microneme protein Etmic-2</i>	<i>Microneme Etmic-2</i>
Emax_0247	0	1	137	0	Proteína precursora de SERPINA	<i>SERPIN, SERine Proteinase INhibitors</i>	<i>Protease inhibitor I4</i>
Emax_0264	90	150	25	0	Proteína putativa não caracterizada de <i>E. tenella</i>	*ND	<i>EGF-like region</i>
Emax_0357	5	15	0	0	Dinamina	<i>Dynamain_M, Dynamain central region</i>	<i>Dynamain central region</i>
Emax_0639	37	59	2	0	Aspartil proteinase (Eimepsina)	<i>Asp, Eukaryotic aspartyl protease</i>	<i>Peptidase aspartic</i>
Emax_0859	0	0	288	0	Proteína hipotética	*ND	*ND
Emax_0955	40	2	1	0	Aminopeptidase	*ND	<i>Peptidase M1, membrane alanine aminopeptidase</i>
Emax_0999	12	4	1	0	Proteína rombóide	<i>Rhomboid, Rhomboid family</i>	<i>Peptidase S54, rhomboid</i>
Emax_1094	171	225	27	0	Transhidrogenase	<i>PNTB, NAD(P) transhydrogenase beta subunit</i>	<i>NAD(P) transhydrogenase</i>
Emax_1166	79	53	3	0	Proteína da família BT1	<i>BT1, BT1 family</i>	<i>Biopterin transport-related protein BT1</i>
Emax_1232	64	39	7	0	Proteína de desenvolvimento de esporozoíto	<i>Ptr, Secreted/periplasmic Zn-dependent peptidases, insulinas.</i>	<i>Peptidase M16, zinc-binding site</i>

*ND- não foram obtidos resultados de Interpro e/ou RPS-BLAST para os transcrito de *Eimeria maxima*

Para os genes diferencialmente expressos selecionados foi realizada uma análise comparativa dos perfis de expressão obtidos por ORESTES e/ou ESTs entre leituras de outras espécies de *Eimeria*, cujo produto era similar aos de *E. maxima* (item 3.19). Ao analisar os resultados (Tabela 18), observamos para a maioria dos perfis de expressão gênica em oocistos são distintos entre as espécies de *Eimeria* e somente um transcrito apresentou perfil semelhante entre as três espécies. Dos transcritos de *E. maxima* analisados, 2 apresentaram perfis similares com o de *E. acervulina* e 3 com os de *E. tenella*. Ainda notamos que cerca de 85% dos genes estudados em *E. maxima* apresentaram uma maior transcrição para o mesmo estágio em pelo menos uma das espécies comparadas.

Tabela 18- Comparação dos perfis de expressão em três fases da esporulação de oocistos dos genes selecionados para estudos de RT-qPCR de *E. maxima* com os transcritos similares de *E. acervulina* e *E. tenella*. Estão apresentados na tabela a identificação dos transcritos de *E. maxima*, *E. tenella* e de *E. acervulina* e os respectivos perfis de expressão. Os estagios evolutivos avaliados foram (Ou) oocistos não esporulados, (Op) oocistos parcialmente esporulados e (Os) oocistos esporulados Os perfis de expressão semelhantes à *E. maxima* estão grafados em azul e os diferentes em laranja.

<i>E. maxima</i>		<i>E. tenella</i>		<i>E. acervulina</i>	
ID do cDNA	Perfil de expressão	ID do cDNA	Perfil de expressão	ID do cDNA	Perfil de expressão
Emax_0017	Os>Ou	Eten_0356	Os	Eace_1254	Os
Emax_0130	Os	*ND	*ND	Eace_0315	Os>Op
Emax_0208	Os>Op	Eten_0776	Os>Op	*ND	*ND
Emax_0247	Os>Op	Eten_3216	Os>Ou -	Eace_0134	Os>Op
Emax_0264	Op>Ou>Os	Eten_1237	Op>Ou>Os	Eace_0293	Ou>Op>Os
Emax_0357	Op>Ou	*ND	*ND	Eace_0170	Ou>Op
Emax_0639	Op>Ou>Os	Eten_2297	Ou>Op>Os	Eace_0021	Os>Ou>Op
Emax_0859	Os	*ND	*ND	Eace_0398	Os>Op
Emax_0955	Ou>Op>Os	Eten_6192	**	Eace_0332	Ou
Emax_0999	Ou>Op>Os	Eten_1120	Op>Os>Ou	Eace_0780	Ou
Emax_1094	Op>Ou>Os	Eten_1199	Op>Ou>Os	Eace_0433	Ou>Op>Os
Emax_1166	Ou>Op>Os	Eten_1535	Ou>Op>Os	Eace_0634	Ou>Op>Os
Emax_1232	Ou>Op>Os	Eten_1291	Op>Ou>Os	Eace_1266	Ou>Op>Os

*ND- não foi encontrada similaridade significativa ao transcrito de *Eimeria maxima*

**Não apresentou contagem nos estágios Ou, Op ou Os.

4.18.3 Expressão diferencial por RT-qPCR

Antes da quantificação da expressão gênica por PCR em tempo real, foram realizados experimentos para cálculo da curva de eficiência dos *primers* e da reação.

A avaliação da eficiência dos *primers* foi realizada empregando um mix de cDNA dos três estágios: oocistos não esporulados (Ou), parcialmente esporulados (Op) e esporulado (Os) (item 3.19) e diferentes concentrações (200nM a 500nM) de *primers*. Para a realização dos experimentos, foi escolhida a concentração mínima dos *primers* para a qual era possível se obter a saturação da reação de PCR (Tabela 19).

Tabela 19- Concentração de *primers* utilizada nos experimentos de PCR em tempo real. Estão apresentadas na tabela a identificação do transcrito de *E.maxima*, a descrição do produto e a concentração dos *primers*.

ID do cDNA	Produto	Concentração dos <i>primers</i>
Emax_0017	Proteína de choque térmico HSP90	400 nM
Emax_0130	Antígeno não transmembranar	300 nM
Emax_0208	Proteína de micronema 2	200 nM
Emax_0247	Proteína precursora de SERPINA	400 nM
Emax_0264	Proteína putativa não caracterizada de <i>E. tenella</i>	500 nM
Emax_0357	Dinamina	300 nM
Emax_0608	Ubiquitina	200 nM
Emax_0639	Aspartil proteinase (Eimepsina)	400 nM
Emax_0859	Proteína hipotética	300 nM
Emax_0955	Aminopectidase	300 nM
Emax_0999	Proteína rombóide	400 nM
Emax_1094	Transhidrogenase	400 nM
Emax_1166	Proteína da família BT1	300 nM
Emax_1232	Proteína de desenvolvimento de esporozoíto	400 nM

Em seguida, foram realizados experimentos para determinação da curva de eficiência da reação a partir de um mix de cDNA dos três estágios estudados (Ou, Op e Os) (item 3.19). Para os diferentes genes estudados, as eficiências das reações foram muito semelhantes entre si, variando de 1,60 a 1,70 (Tabela 20).

Tabela 20- Eficiência da reação obtida a partir das reações realizadas.

ID do cDNA	Produto	Eficiência da reação
Emax_0017	Proteína de choque térmico HSP90	1,69
Emax_0130	Antígeno não transmembranar	1,64
Emax_0208	Proteína de micronema 2	1,67
Emax_0247	Proteína precursora de SERPINA	1,60
Emax_0264	Proteína putativa não caracterizada de <i>E. tenella</i>	1,63
Emax_0357	Dinamina	1,64
Emax_0608	Ubiquitina	1,64
Emax_0639	Aspartil proteinase (Eimepsina)	1,64
Emax_0859	Proteína hipotética	1,71
Emax_0955	Aminopeptidase	1,71
Emax_0999	Proteína rombóide	1,64
Emax_1094	Transhidrogenase	1,63
Emax_1166	Proteína da família BT1	1,63
Emax_1232	Proteína de desenvolvimento de esporozoíto	1,63

A quantificação da expressão relativa dos 13 genes selecionados para os estágios oocistos não esporulados, parcialmente esporulados e esporulados está mostrada na Figura 30.

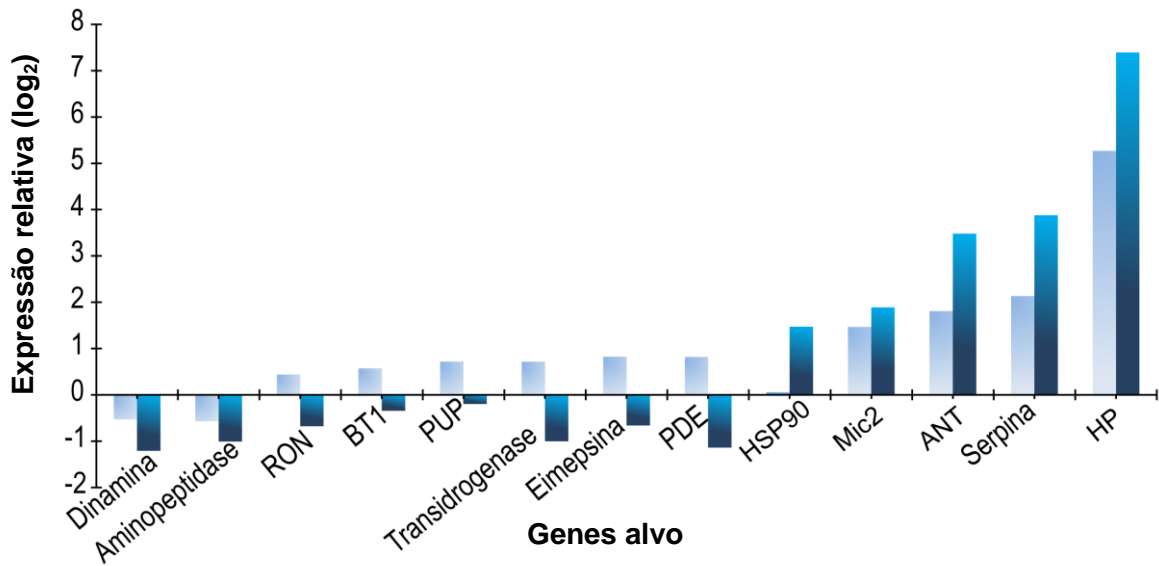


Figura 30: Expressão relativa (\log_2) obtida pela técnica de Real-time PCR (qRT-PCR) para os genes codificantes de Dinamina; Aminopeptidase; RON (proteína rombóide); Proteína da família BT1; PUP (proteína putativa não caracterizada); Transhidrogenase; Eimepsina; Proteína de desenvolvimento de esporozoíto (PDE); HSP 90 (proteína de choque termico); Mic2 (proteína de micronema 2); ANT (antígeno não transmembranar); Serpina e HP (proteína hipotética) em oocistos parcialmente esporulados e esporulados em relação aos oocistos não esporulados (nível basal, eixo x). Oocistos parcialmente esporulados (■), oocistos esporulados (■).

Ao analisar a expressão gênica dos 13 genes escolhidos, observa-se que o perfil transcricional é diferente entre os estágios avaliados, de fato, 15% dos genes são mais transcritos em oocistos não esporulados (Ou), 39% em oocistos esporulados (Os) e 46% dos genes em oocistos parcialmente esporulados (Op).

Dentre o grupo dos transcritos mais expressos em Ou estão os genes codificantes para a dinamina e aminopeptidase. Neste grupo, a transcrição gênica tende a decrescer durante a esporulação (Ou>Op>Os). Tanto para a dinamina quanto para a aminopeptidase, a expressão gênica nos oocistos não esporulados é cerca de 2 vezes maior do que em oocistos esporulados.

Há um segundo grupo onde os transcritos são preferencialmente observados em oocistos parcialmente esporulados, seguido de oocistos não esporulados e oocistos esporulados (Op>Ou>Os). Dentro deste estão a proteína putativa não caracterizada (PUP), eimepsina, proteína rombóide (RON), transhidrogenase, proteína da família BT1 e proteína de desenvolvimento do esporozoíto (PDE). Para a transhidrogenase, eimepsina e proteína de desenvolvimento do esporozoíto (PDE) a

transcrição é cerca de 3 a 4 vezes maior em Op do que em Os. Para as demais, a transcrição em Op é aproximadamente 2 vezes maior que em Os.

Diferentemente destes dois grupos onde Os foi o estágio de menor transcrição dos genes avaliados, no terceiro conjunto Os foi o estágio que apresentou maior nível de expressão. Neste grupo a expressão gênica tende a aumentar no decorrer da esporulação ($Ou < Op < Os$), este padrão foi verificado para os genes codificantes de proteína de choque térmico (HSP90), antígeno transmembranar (ANT), proteína micronema 2 (Mic2), serpina e proteína hipotética (HP). Grandes variações de níveis de expressão relativa foram observadas entre os genes deste grupo. Enquanto que em HSP90 e proteína de micronema 2 a transcrição gênica em Os é respectivamente 2,7 e 3,7 vezes maior que em Ou, para a transidrogenase e serpina a expressão em Os é respectivamente 11,15 e 14,7 vezes maior que em Ou. Já para a proteína hipotética, a transcrição gênica em Os é 168 vezes maior que em Ou e em Op e é 38 vezes maior que em Ou.

Estudos comparativos dos perfis de expressão gênica obtido pelas técnicas RT-qPCR (RT- Real time PCR) e ORESTES de *Eimeria maxima* foram realizados para os 13 genes estudados (Tabela 21). Ao analisar em conjunto os estágios que apresentaram maior e menor transcrição gênica observamos que para ambas as técnicas, o perfil de expressão foi concordante para 62% dos genes estudados. Se compararmos somente os estágios de maior transcrição gênica, observa-se que os perfis são concordantes para 70% dos genes. Para os estágios de menor transcrição, os resultados foram concordantes para 92 % dos genes (Tabela 21).

Tabela 21 – Comparação da expressão gênica obtida pelas técnicas de qRT-PCR (RT- Real time PCR) e ORESTES para as fases da esporulação de *E.maxima* de maior e menor transcrição de 13 genes estudados. Estão descritos na tabela o produto e a fase da esporulação de maior e menor transcrição para as técnicas de qRT-PCR e ORESTES. Oocistos não esporulados (Ou), oocistos parcialmente esporulados (Op) e oocistos esporulados (Os).

Produto	Fase de maior transcrição		Fase de menor transcrição	
	RT-qPCR	ORESTES	RT-qPCR	ORESTES
Proteína de choque térmico HSP90	Os	Os	Ou	Op
Antígeno não transmembranar	Os	Os	Ou	Ou/Op
Proteína de micronema 2	Os	Os	Ou	Ou
Proteína precursora de SERPINA	Os	Os	Ou	Ou
Proteína putativa não caracterizada	Op	Op	Os	Os
Dinamina	Ou	Os	Os	Os
Aspartil proteinase (Eimepsina)	Op	Op	Os	Os
Proteína hipotética	Os	Os	Ou	Ou/Op
Aminopeptidase	Ou	Os	Os	Os
Proteína rombóide	Op	Ou	Os	Os
Transhidrogenase	Op	Op	Os	Os
Proteína da família BT1	Op	Ou	Os	Os
Proteína de desenvolvimento de esporozoíto	Op	Ou	Os	Os

5 DISCUSSÃO

Dentre as espécies de *Eimeria* que infectam a galinha doméstica, a *Eimeria maxima* é uma das mais patogênicas, é comumente observada no campo e tem larga distribuição mundial (Martin et al., 1997; Schnitzler et al., 1998; Schnitzler et al., 1999; Lew et al., 2003). Apesar da grande importância, estudos da análise da expressão gênica nesta espécie são bastante escassos. Até o momento, existem apenas três trabalhos sobre o perfil transcricional em *E. maxima* (Basak et al., 2006; Schwarz et al., 2010; Dong et al., *in press*) e nenhum destes é sobre o processo de esporulação.

Um dos estágios de grande importância para o ciclo de vida de coccídias é o oocisto que uma vez no ambiente, torna-se infectante após a esporulação ou esporogonia. A esporogonia envolve várias etapas, dentre as quais, três divisões nucleares (uma meiose e duas mitoses) (Vetterling e Doran, 1969; Hammond, 1973; Ryley, 1973), culminando na formação do esporozoíto. Durante este processo, também há metabolismo de vários compostos, incluindo carboidratos (Edgar et al., 1944; Wilson e Fairbairn, 1961; Ryley et al., 1969; Ryley, 1973; Wang et al., 1975; Joyner, 1982; Wang, 1982) e lipídeos (Wilson e Fairbairn, 1961; Joyner, 1982). Vários mecanismos moleculares devem estar envolvidos neste processo que ainda é pouco conhecido e estudado.

Grande parte dos estudos sobre expressão gênica em oocistos foi realizado para genes específicos, principalmente de *E. tenella*, *E. acervulina* e *E. bovis*. Para *E. maxima* há poucos artigos sobre este assunto e os disponíveis estudaram o gene codificante para TFP em oocistos (Smith et al., 1994; Witcombe et al., 2003).

Há poucos trabalhos sobre a expressão gênica de larga escala em oocistos de *Eimeria* spp. (Abrahamsen et al., 1993; Miska et al., 2004; Basak et al., 2006; Han et al., 2007; Han et al., 2010; Dong et al., *in press*). Estes estudos avaliaram a transcrição gênica em oocistos não esporulados e/ou esporulados, não analisando a expressão nas fases intermediárias da esporulação.

Considerando a importância econômica, a ocorrência no campo, a escassez de dados de expressão gênica e do perfil transcricional da esporulação em *E. maxima*, decidimos estudar a expressão gênica diferencial em três fases da esporulação de oocistos de *E. maxima*. Para tanto, empregamos a metodologia ORESTES que dentre outras vantagens, é capaz de gerar um perfil qualitativo normalizado da expressão gênica (Dias Neto et al., 2000).

Desta forma, as leituras ORESTES foram obtidas a partir de bibliotecas de cDNA de diferentes fases da esporulação de oocistos de *E. maxima* (oocistos não esporulados, parcialmente esporulados e esporulados).

5.1 Purificação dos Parasitas e extração de RNAm

Após a infecção das aves e purificação dos oocistos, o rendimento médio obtido foi de 2.000 oocistos/ave para cada oocisto ingerido totalizando em média 10^7 oocistos/ave. A taxa média de esporulação foi de 75% após 72 horas de incubação a 28 °C. Estes resultados são semelhantes aos relatados anteriormente pelo grupo (Fernandez, 1999). Os resultados das infecções foram satisfatórios, os oocistos empregados foram obtidos em quantidade suficiente e também foi verificado que não estavam contaminados com nenhuma outra espécie de *Eimeria*.

Para a obtenção de oocistos parcialmente esporulados (fase de esporoblasto), foi realizada uma cinética de esporulação. A maior parte (70%) dos esporoblastos foi observada entre 22 e 24 horas do início da esporulação. Em *E. maxima*, os oocistos na fase de esporoblastos podem ser visualizados a partir de 16 horas da esporulação (Canning e Anwar, 1968), entretanto não há dados na literatura que relatem o momento no qual a maioria destes oocistos é visualizada.

Os RNAs mensageiros de oocistos não esporulados, parcialmente esporulados e esporulados foram extraídos e avaliados quanto à integridade e quantidade. Todos os RNAs mensageiros estavam íntegros e isentos de contaminação por DNA genômico de *E. maxima* e foram utilizados para a etapa de construção das bibliotecas de ORESTES.

5.2 Construção de bibliotecas de ORESTES

Para a obtenção das bibliotecas de cDNA foi empregada a técnica ORESTES (Dias Neto et al., 2000) que é uma modificação da estratégia convencional de

obtenção de ESTs, uma vez que as sequências são produzidas ao longo dos transcritos, ao invés de somente a partir das extremidades dos cDNAs. Esta técnica consiste na geração de moldes de cDNA por RT-PCR sob condições de baixa estringência, usando *primers* arbitrários (Dias Neto et al., 2000). Além da geração de novos transcritos, esta técnica pode auxiliar na identificação de introns e exons para a anotação de genomas (Mello et al., 2009) e permitir a identificação de transcritos raros (Camargo et al., 2001).

Até o momento, não há dados de expressão gênica para parasitas do Filo Apicomplexa utilizando a metodologia ORESTES. No presente trabalho foi gerado um total de 20.149 sequências do tipo ORESTES, das quais, 4.424 eram de oocistos não esporulados, 3.480 de oocistos parcialmente esporulados e 12.085 de oocistos esporulados.

Com o intuito de eliminar sequências de baixa qualidade e sequências contaminantes, as leituras geradas foram filtradas, conforme descrito no item 3.9, aceitando-se 75% das sequências geradas. Resultado semelhante foi relatado em um estudo do transcriptoma de *Schistosoma mansoni* utilizando a metodologia ORESTES, onde foi aceita aproximadamente 76% das sequências após a filtragem (Verjovski-Almeida et al., 2003). Em um estudo do transcriptoma empregando ORESTES, a partir de células de tumor de mama, cerca de 79% das sequências geradas foram aceitas (Dias Neto et al., 2000). De fato, sabe-se que bibliotecas de ESTs, incluindo ORESTES podem conter sequências contaminantes (Nagaraj et al., 2007) e de baixa qualidade. Desta forma, a taxa de aceitação das sequências está dentro dos padrões esperados para a técnica empregada.

O processo de agrupamento das sequências é um passo importante para a análise dos transcritos. Nesta etapa, sequências semelhantes são agrupadas e assim consideradas como pertencentes a um mesmo gene. Antes de realizar o agrupamento das sequências é necessário padronizar os melhores parâmetros a serem utilizados para a reconstrução dos transcritos. Quando se emprega baixa estringência pode ocorrer junção de sequências parálogas ou, pior ainda, de sequências não relacionadas. Por outro lado, se a estringência é alta demais, há uma maior fragmentação da informação, resultando em sequências consenso mais curtas. Ferro (2008) testou vários parâmetros para o agrupamento das sequências ORESTES de *Eimeria tenella* no programa CAP3, sendo escolhido p 90 e y 250, parâmetros estes também utilizados no presente trabalho (item 3.10). Após o

agrupamento das sequências foi obtido um total de 3.298 eventos, sendo 1.207 *contigs* e 2.091 *singlets*.

Em *Toxoplasma gondii*, protozoários da classe Coccidia, o tamanho do genoma é de 63 Mb e estima-se que tenha 7800 genes codificantes (Mullapudi et al., 2009; Weiss et al., 2009). Baseando na densidade gênica de *Toxoplasma gondii*, estima-se que *Eimeria maxima* deva ter entre 6800 a 7400 genes codificantes, já que o tamanho de seu genoma é de 55 a 60 Mb. Partindo desta premissa, ao analisar os resultados da montagem dos transcritos ORESTES, acreditamos que apenas uma fração do transcriptoma de *E. maxima* deve ter sido coberta, uma vez que foram obtidos 3.298 eventos. Possíveis razões para esta baixa cobertura podem estar relacionadas à geração insuficiente de leituras, expressão de poucos genes em oocistos nas diferentes fases da esporulação, e/ou regulação transcricional específica para cada fase da esporulação.

Alta redundância (cerca de 78%) foi observada após a montagem das bibliotecas de ORESTES de *E. maxima*. Estudos do transcriptoma de *Schistosoma mansoni* (Verjovski-Almeida et al., 2003), *Apis mellifera* (Nunes et al., 2004) e *Trypanosoma rangeli* (Grisard et al., 2010) empregando a metodologia ORESTES, também obtiveram taxas de redundância de 75%, 73% e 60% respectivamente. Redundância alta também foi observada na montagem das sequências ORESTES de vários estágios do ciclo biológico de *Eimeria tenella* e de *Eimeria acervulina*, ambas realizadas pelo nosso grupo.

Como a alta redundância foi observada desde a geração das primeiras leituras de ORESTES de *E. maxima* (dados não mostrados), este resultado pode ser um indicativo de que um pequeno conjunto de genes deve ser altamente expresso em cada fase da esporulação.

O tamanho médio dos *contigs* obtidos foi de 558,32 nucleotídeos, sendo que muitos destes (44%) apresentavam tamanho entre 300 e 600 nucleotídeos. Resultados semelhantes foram obtidos em transcriptoma de *Apis mellifera* (Nunes et al., 2004) e de epimastigota e tripomastigota de *Trypanosoma rangeli* (Grisard et al., 2010) nos quais após o agrupamento das sequências ORESTES, foram observados *contigs* de tamanho médio de 519 e 370 nucleotídeos, respectivamente.

Analisando a composição dos *contigs* foi observado que a maioria (85,5%) possuía um número pequeno de leituras (<10) e uma minoria (1,66%) apresentava mais do que 100 leituras. Resultados semelhantes foram observados em bibliotecas

ORESTES de *Apis melífera* (Nunes et al., 2004), de *Biomphilaria glabrata* (Lockyer et al., 2007; Hanelt et al., 2008) e de queratinócitos granulares humanos (Toulza et al., 2007).

Os nossos dados sugerem que um pequeno conjunto de genes poderia ser altamente transcrito e que a grande parte dos genes é pouco expressa nos diferentes estágios da esporulação. Dados semelhantes foram observado em um estudo da expressão gênica em merozoítos e esporozoítos de *Eimeria tenella*, utilizando a metodologia SAGE no qual, grande parte (66%) das *tags* geradas eram únicas e somente 1% dos reconstruídos possuía mais do que 50 *tags* (Novaes, 2009).

É descrito que os transcritos obtidos pela técnica ORESTES são gerados preferencialmente a partir das porções centrais dos genes os quais representam as regiões codificadoras (Dias Neto et al., 2000). Para verificar a localização dos transcritos gerados, o cDNAs reconstruídos foram mapeados nas sequências completas de cDNA de *E. maxima* depositadas no banco de dados Genbank. Foi verificado que cerca de 70% dos transcritos selecionados mapearam na região codificante das sequências analisadas. Apesar da quantidade reduzida de sequências completas de cDNA de *E. maxima* disponíveis no GenBank, o percentual de transcritos que mapearam na região codificante dos genes foi similar ao relatado em demais trabalhos que empregaram ORESTES para a construção de bibliotecas de cDNA de humano (Dias Neto et al., 2000; Camargo et al., 2001; Sakabe et al., 2003) e de boi (da Mota et al., 2004). Apesar do genoma de *Eimeria maxima* estar disponível, não foi possível identificar se os transcritos mapeavam nas regiões codificadoras dos genes, pois no momento da realização da análise, o genoma ainda não estava predito e anotado.

Para o mapeamento dos *contigs* no genoma de *E. maxima*, foi empregada uma abordagem simples, na qual, foram realizadas buscas de similaridade entre os transcritos reconstruídos e os *contigs* do genoma utilizando o programa BLASTN. Para tanto, diferentes percentuais de similaridade e de alinhamento entre os transcritos ORESTES e os *contigs* do genoma foram testados. Cerca de 77% dos *contigs* mapearam no genoma com uma similaridade acima de 95% e tamanho mínimo de alinhamento de 40%, ou seja, o tamanho do alinhamento entre o transcrito e o genoma representava pelo menos 40% do tamanho do *contig*. Resultado semelhante foi relatado num estudo de expressão gênica de diversos

tumores humanos utilizando a técnica ORESTES, no qual cerca de 88% dos genes puderam ser mapeados no genoma humano (Mello et al., 2009). Como este mapeamento não foi realizado para o genoma predito de *E. maxima*, consideramos que a taxa de mapeamento obtida foi bastante satisfatória.

5.3 Anotação automática das sequências

Após o agrupamento, os transcritos reconstruídos foram submetidos a um protocolo de anotação automática desenvolvida pelo grupo de bioinformática do laboratório. A anotação de genes foi obtida a partir de dados de similaridade com sequências biológicas de função conhecida, encontro de motivos proteicos, regiões transmembranares, peptídeo sinal, entre outros. Dentre as evidências utilizadas para a anotação automática, a busca de similaridade utilizando BLAST (Altschul et al., 1997) constituiu uma das principais etapas de caracterização das sequências, especialmente quanto ao seu grau de conservação e sua possível função, sendo a principal abordagem escolhida para análise dos transcritos.

Dos 1207 *contigs* totais, somente 28% dos cDNAs reconstruídos apresentaram resultados positivos de BLAST contra a base de dados nr. A partir destes dados, podemos concluir que a maior parte dos transcritos gerados provavelmente ainda não foi caracterizada ou mesmo descrita, e isto pode estar relacionado ao fato destas sequências ainda não terem sido identificadas, ou ainda, serem específicas dos oocistos ou da espécie em questão. Estudos de expressão gênica em oocistos de *E. maxima* também mostraram que a maioria (81%) dos transcritos gerados não apresentava similaridade significativa com proteínas descritas (Dong et al., *in press*). Para *E. tenella*, 64% das proteínas de oocistos obtidas por hibridização subtrativa foram classificadas como hipotéticas (Miska et al., 2004). Mesmo para merozoítos e esporozoítos de *Eimeria* spp., que são os estágios mais estudados, grande quantidade de proteínas hipotéticas também foi observada em bibliotecas de ESTs (Abrahamsen et al., 1995; Wan et al., 1999; Ng et al., 2002; Refega et al., 2003; Schwarz et al., 2010) e de SAGE (Novaes, 2009).

Proteínas hipotéticas também foram frequentes em estudos de transcriptoma de demais parasitas, como os do Filo Apicomplexa. Em *Plasmodium falciparum*,

cerca de 60% das proteínas preditas no genoma não apresentaram similaridade com proteínas de outros organismos (Gardner et al., 2002). Em *Toxoplasma gondii* metade dos genes preditos está anotado como proteínas hipotéticas (Xia et al., 2008). Após análise de busca de similaridade de ESTs de diferentes parasitas, como *Plasmodium falciparum*, *Toxoplasma gondii*, *Eimeria tenella*, *Sarcocystis neurona* e *Neospora caninum*, com o banco de dados nr (Genbank) foi observado que entre 51 a 74% das proteínas eram hipotéticas (Li et al., 2003).

A grande quantidade de proteínas hipotéticas observada no nosso estudo indica que muitos genes ainda não são conhecidos, sendo necessários diversos estudos complementares de expressão proteica, localização e de atividade para inferir possíveis funções destas proteínas.

Dos 329 *contigs* que apresentaram similaridade significativa a proteínas já descritas, a maioria (74,2%) era similar a proteínas de parasitas do Filo Apicomplexa, o que é um bom indício tanto para a caracterização destes reconstruídos, quanto para a credibilidade do protocolo de anotação utilizado. Destas, 19% eram semelhantes a proteínas descritas em parasitas do gênero *Eimeria*. Algumas destas proteínas, tais como, proteína de micronema, serpina, aminopeptidase, proteína de choque térmico, proteína quinase já foram descritas em estudos do transcriptoma de oocistos de *Eimeria tenella* (Miska et al., 2004) e *Eimeria maxima* (Han et al., 2010; Dong et al., *in press*).

Analisando-se os dados de ontologia gênica (GO), apenas 20,3% das sequências dos *contigs* puderam ser classificadas em uma ou mais ontologias principais (função molecular, processo biológico e componente celular). Ao analisar separadamente cada uma das três ontologias gênicas principais, observamos que para processo biológico, grande parte dos produtos (10%) está relacionada a processos metabólicos, particularmente, processos biossintéticos. Para o componente celular, há uma maior porcentagem de sequências (6,20%) relacionadas à célula, sendo as categorias de componentes intracelulares e do núcleo as mais frequentes. Para a função molecular, a categoria com o maior número de representantes estava relacionada à atividade catalítica, incluindo atividade de transferase e hidrolase. Os termos de GO apresentaram boa correlação com os dados de anotação, a maioria dos transcritos que apresentava resultados positivos por BLAST era relacionada ao metabolismo. De fato, sabe-se que durante a esporulação, vários compostos são metabolizados, tais como, carboidratos e

lipídeos (Edgar et al., 1944; Wilson e Fairbairn, 1961; Ryley et al., 1969; Ryley, 1973; Wang et al., 1975; Joyner, 1982; Wang, 1982).

5.4 Expressão gênica diferencial

Para identificar o perfil de expressão gênica nas três fases da esporulação de oocistos foi empregado o *Northern* digital (Audic e Claverie, 1997), que é uma abordagem computacional que possibilita a identificação de genes diferencialmente expressos em bibliotecas de larga escala (SAGE, MPSS, ESTs). Esta técnica baseia-se no fato de que cada cDNA reconstruído pode estar representado por sequências de diferentes estágios evolutivos e a abundância destas pode refletir o nível relativo de transcrição dos respectivos cDNAs, permitindo assim a identificação de genes diferencialmente expressos.

A análise dos reconstruídos por *Northern* digital mostrou que apenas 27% dos transcritos eram compartilhados entre dois ou mais estágios estudados, mostrando um perfil de expressão diferencial entre as fases da esporulação. Miska et al., (2004) também observaram que uma pequena porcentagem (6,7%) dos transcritos era compartilhada entre oocistos esporulados e não esporulados de *Eimeria tenella*. Ellis e Thurlby (1991) demonstraram que em oocistos de *E. maxima*, há genes que são diferencialmente transcritos durante a esporulação.

Expressão gênica estágio-específica também foi relatada em outros parasitas do Filo Apicomplexa, tais como: *Toxoplasma gondii* (Li et al., 2003; Gopalakrishnan e Lopez-Estrano, 2010; Lopez-Estrano, 2010), *Neospora caninum* (Kang et al., 2008a; Kang et al., 2008b), *Babesia bovis* (de Vries et al., 2006), *Plasmodium falciparum*. (Bozdech et al., 2003; Gopalakrishnan e Lopez-Estrano, 2010; Lopez-Estrano, 2010).

Como os nossos dados mostravam que a expressão gênica era diferencial entre os estágios estudados, o próximo passo foi avaliar se estes perfis de expressão gênica eram correlacionados entre si, e ainda, se as relações gênicas faziam sentido em termos da biologia do parasita. Para gerar um agrupamento hierárquico dos genes e calcular as distâncias entre os clados obtidos, os resultados

de expressão gênica das três fases da esporulação de oocistos, adicionados de leituras de merozoítos de *E. maxima* (item 3.15) foram agrupados pelo programa Simcluster (Vencio et al., 2007). Após a análise, foi verificado que independentemente do método de distância utilizado, as árvores obtidas apresentavam topologia semelhante (dados não mostrados).

O fato dos oocistos não esporulados, parcialmente esporulados e esporulados agruparem num único clado e os merozoítos num clado externo, mostra que os oocistos estão mais correlacionados do que os merozoítos, conforme esperado para a biologia do parasita. Analisando o clado dos oocistos, observou-se que os oocistos não esporulados e parcialmente esporulados estão mais relacionados do que oocistos esporulados. Leituras provenientes de vários estágios de *E. acervulina* (ORESTES) e de *E. tenella* (ESTs/ORESTES) quando agrupadas, também mostraram árvores de topologia semelhante à obtida para *E. maxima* (manuscrito em preparação). Interessante notar que apesar da esporulação não ser um processo sincrônico (Vetterling e Doran, 1969; Hammond, 1973; Ryley, 1973), o padrão da expressão gênica foi semelhante entre as três espécies.

Apesar da análise digital dos transcritos não ser recomendada para estudos de expressão diferencial em bibliotecas normalizadas (Ewing et al., 1999), os resultados do agrupamento hierárquico dos transcritos obtidos foram consistentes e condizentes com a biologia do parasita, indicando que para este estudo, o *Northern* digital teve boa aplicabilidade.

Para determinação e classificação dos genes diferencialmente expressos foi empregado o programa Kemp (Varuzza, 2008; Varuzza e Pereira, 2010). O Kemp é um programa baseado em um teste frequentista desenvolvido para a comparação de perfis de expressão digital sendo empregado em estudos de transcriptoma, tais como, análise da expressão gênica em carcinoma de células escamosas da laringe (Silveira et al., 2008) e análise seriada da expressão gênica de merozoítos e esporozoítos de *E. tenella* (Novaes, 2009). Este teste calcula os valores de p (*p-value*) que representam uma medida numérica da coerência dos dados observados em relação à hipótese nula (não diferenciação da expressão dos transcritos). Para o cálculo dos valores de p, este teste utiliza uma técnica exata independente da quantidade de bibliotecas e das contagens totais dos transcritos em todas as bibliotecas, mas altamente dependente das verossimilhanças (função de probabilidade condicional). Para a seleção de genes diferencialmente expressos,

normalmente um único valor de corte canônico é utilizado para todos os transcritos ($p \leq 0,1$; 0,05 ou 0,01). Estes valores de corte são utilizados para definir se um transcrito é diferencialmente expresso ou não. Em qualquer perfil de expressão gênica, a abundância relativa de cada transcrito pode variar de forma significativa e a utilização de um nível de corte fixo pode alterar a classificação dos dados de expressão, principalmente para transcritos de baixa abundância. O Kemp possui a grande vantagem de calcular um valor de nível crítico para cada transcrito em particular (α), o que minimiza a combinação linear dos erros do tipo I (falso positivo) e II (falso negativo). Como nosso objetivo foi de analisar os transcritos diferencialmente expressos, foi utilizado um peso maior (4) para a hipótese nula (erros do tipo I) selecionando assim os transcritos realmente diferencialmente expressos, minimizando os erros do tipo II (falso negativo). A classificação dos genes diferencialmente expressos foi realizada de forma automática. Os *contigs* com valores de p menor do que α (nível crítico) foram classificados como diferencialmente expressos (Varuzza, 2008).

A maioria (68,6%) dos *contigs* foi classificada como não diferencialmente expressos, dos quais 76% eram proteínas hipotéticas e apenas 17% apresentavam similaridade significativa com genes de função conhecida.

Os *contigs* classificados como diferencialmente expressos apresentavam diferentes perfis de expressão, a maioria (73%) tinha contagem em apenas um estágio e apenas uma pequena parte (12,4%) era compartilhada entre os estágios estudados. De fato, a figura do *Heat map* mostra este padrão diferencial da expressão gênica. Nota-se que distintos conjuntos de genes são expressos em cada fase da esporulação (oocisto não esporulado, parcialmente esporulado e esporulado) indicando que os genes são diferencialmente transcritos durante este processo, ou ainda, que poderia estar ocorrendo regulação transcricional específica nas fases da esporulação estudadas.

A regulação transcricional é um importante mecanismo para o desenvolvimento, homeostase e adaptação das células ao meio ambiente. Entretanto até o momento, esses mecanismos ainda não são totalmente compreendidos em protozoários. Há evidências de que muitas das maquinarias empregadas por eucariotos superiores para o controle da expressão gênica em níveis transcricionais e pós transcricionais também sejam utilizadas em protozoários (Gomez et al., 2010). Regulação transcricional já foi descrita em parasitas do Filo

Apicomplexa, como por exemplo, *Plasmodium falciparum* (Gissot et al., 2005; Gomez et al., 2010) e *Toxoplasma gondii* (Mullapudi et al., 2009; Weiss et al., 2009). Nestes parasitas, a transcrição da maioria dos genes só ocorre quando há demanda fisiológica, ou seja, os genes só são ativados quando a função biológica torna-se necessária.

Diversos trabalhos têm demonstrado que em *Eimeria* spp. genes codificadores de várias proteínas são diferencialmente transcritos em oocistos em distintas fases da esporulação, tais como: MOP, proteína de oocisto (Fetterer e Barfield, 2003); eimepsina (Laurent et al., 1993; Jean et al., 2001a; Miska et al., 2004); proteína quinase dependente de ciclina de *E. tenella*, EtCRK2 (Kinnaird et al., 2004); aminopeptidase N (Miska et al., 2004; Fetterer et al., 2005); serino proteases (Fetterer et al., 2007b); antígeno SO7 (Fetterer et al., 2007a); proteína associada ao corpo refrátil (Abrahamsen et al., 1994); proteína de choque térmico 90, HSP90 (Clark et al., 1996; Miska et al., 2005; Péroval et al., 2006); proteínas de micronema (Ryan et al., 2000; Miska et al., 2004); proteínas rombóides; prolil endopeptidase; transhidrogenase; proteína quinase, (Miska et al., 2004); manitol-1-fosfato-desidrogenase (Schmatz, 1997; Allocco et al., 1999); fator depolimerizante de actina; actina (Xu et al., 2008) e lactato desidrogenase (Schaap et al., 2004).

Os genes expressos nos diferentes estágios da esporulação foram analisados com o intuito de verificar se havia genes de funções correlatas dentro de cada fase estudada. Infelizmente, não foi possível fazer esta categorização uma vez que muitas proteínas eram hipotéticas. De fato, a maioria (69%) dos transcritos diferencialmente expressos foi classificada como proteínas hipotéticas e apenas 21% dos produtos eram similares a proteínas de função conhecida.

Com a finalidade de estudar a transcrição diferencial entre as três fases da esporulação, 100 transcritos diferencialmente expressos mais abundantes foram selecionados e estudados. Além da anotação automática, as proteínas classificadas como hipotéticas que estavam presentes dentre estes transcritos, foram também analisadas pelo programa Threader (Jones, 1998), com o intuito de identificar possíveis funções a partir de estruturas proteicas conservadas. Um total de 71% das proteínas apresentou enovelamento predito similar com outras proteínas depositadas no banco de dados RCSB PDB (*The Protein Data Bank*). Verificou-se a existência de proteínas de diversas funções, muitas das quais, relacionadas ao metabolismo. A partir destes dados, estudos complementares poderão ser

realizados para um melhor esclarecimento da função e importância destas proteínas não só na esporulação, como também, para demais estágios do ciclo evolutivo de *Eimeria* spp.

Quando os 100 transcritos selecionados foram comparados às sequências ORESTES/ESTs de *E. tenella* e ORESTES de *E. acervulina* (item 3.18) foi observado que 56% destes eram similares aos reconstruídos de *E. acervulina* e 59% aos de *E. tenella*. Em relação ao perfil de expressão gênica em oocistos, somente 18% e 22% destes transcritos apresentavam o mesmo perfil em *E. acervulina* e em *E. tenella*, respectivamente. Portanto, a despeito de serem genes ortólogos, o perfil de expressão gênica nas fases de oocistos é diferencial entre as espécies, indicando que a regulação da transcrição destes genes pode ser espécie-específica. Por outro lado, como a análise comparativa do perfil de expressão gênica entre as três espécies de *Eimeria* foi baseada em produtos proteicos traduzidos que apresentavam similaridade entre si pelo programa BLAST, não significa necessariamente que se referem aos mesmos genes. Estudos experimentais são necessários para confirmar estes achados.

Expressão diferencial entre espécies de *Eimeria* já foi observada para outros genes, como o codificante da proteína de choque térmico (HSP90). Em *E. tenella* a transcrição deste gene é constitutiva, sendo observada em todas as fases estudadas (oocistos nas diferentes fases da esporulação, esporozoítos e merozoítos), enquanto que em *E. acervulina*, a transcrição do mesmo só é detectada em oocistos não esporulados, esporozoítos e merozoítos (Miska et al., 2005).

Com o objetivo de validar os resultados de expressão gênica, a quantificação da transcrição nas diferentes fases da esporulação foi avaliada para um pequeno conjunto de genes. Para tanto, foi utilizada a técnica de PCR em tempo real, método padrão para mensurar a expressão gênica diferencial devido a sua alta especificidade e sensibilidade (Deepak et al., 2007). Diversos trabalhos que estudaram a expressão gênica pela geração de EST/ORESTES utilizaram o RT-qPCR para a validação dos resultados obtidos (Miska et al., 2004; Toulza et al., 2007; Dong et al., *in press*).

Devido ao baixo custo e relativa facilidade de se delinear os experimentos, optamos pela utilização do SYBR[®] Green. Como este corante não é específico e pode gerar sinais falso-positivos ao detectar dímeros de *primers* ou produtos inespecíficos (Ponchel, 2006), empregamos *primers* cuidadosamente desenhados e

analisamos as curvas de *melting*. Vale salientar que o kit utilizado possui uma DNA polimerase *hot-start*, reduzindo a chance de amplificação de produtos indesejáveis.

Para realizar estes experimentos, optamos pela quantificação relativa, que em teoria é mais simples e menos suscetível a erros metodológicos (Pfaffl, 2006). Para a realização dos experimentos, somente foram utilizadas amostras de RNAm de alta qualidade, tratadas com DNase e quantificadas de forma acurada com um kit específico que somente reporta a quantidade de moléculas de ácido nucléico (RNA). A utilização de RNAm apesar de mais cara é recomendada, pois já foi relatado que a proporção entre a quantidade de RNAm e RNAr em amostras de RNA total pode variar em diferentes condições e/ou tecidos, podendo alterar a mensuração da expressão gênica (Morse et al., 2005).

O primeiro desafio foi a busca e a validação de genes constitutivamente expressos (Dheda et al., 2004). Sabe-se que os organismos do Filo Apicomplexa possuem uma expressão gênica altamente estágio específica. Por exemplo, em *Plasmodium falciparum*, o perfil transcricional assemelha-se a uma cascata de expressão gênica que resulta no processo “*just in time*” onde os produtos são transcritos de acordo com a necessidade do parasita (Gissot et al., 2005), o mesmo fenômeno acontece em *Toxoplasma gondii* (Mullapudi et al., 2009). Devido isso, estima-se que nestes parasitas, a quantidade de genes expressos de maneira constitutiva seja baixa (Llinas e DeRisi, 2004).

Há poucos trabalhos que relatam a utilização de genes constitutivamente expressos para realização de RT-qPCR em estudos de expressão gênica em *Eimeria* spp. Alguns artigos empregam genes ribossômicos (Miska et al., 2004; Dong et al., *in press*), o que não seria possível no presente trabalho, já que não extraímos RNA total e sim RNAm. O gene actina também foi utilizado como gene constitutivo (Xu et al., 2008), entretanto, os nossos dados mostram que este gene é diferencialmente expresso nas diferentes fases da esporulação (tabela 9). Novaes (2009) também observou transcrição diferencial deste gene entre merozoítos e esporozoítos de *E. tenella*.

Para seleção dos candidatos, baseamos nossa busca nos resultados de ORESTES utilizando o critério de *p-value* igual ou próximo de 1 e níveis de expressão gênica semelhantes entre as três fases da esporulação. Como o intuito era buscar genes de mesmo perfil de expressão, não se levou em conta se os genes escolhidos apresentavam resultados positivos de BLAST. Oito candidatos foram

estudados: 6 proteínas hipotéticas, uma ubiquitina e uma proteína associada à proliferação celular.

O complexo ubiquitina/proteassoma pertence a uma via regulatória pós-transcricional que regula ampla gama de atividades celulares incluindo reparo de DNA, transcrição, divisão celular, endocitose, tráfico intracelular, entre outros. A ubiquitina apresenta 76 aminoácidos e é extremamente conservada entre os organismos eucarióticos, apresentando taxas de similaridades entre as sequências de humanos, leveduras e parasitas superiores a 98%.

Apenas o gene relacionado à ubiquitina foi constitutivamente expresso, apresentando uma variação de Ct <0,5. Expressão de ubiquitina de forma constitutiva tem sido demonstrada em vários organismos. Em determinados tecidos de humanos, o gene da ubiquitina é constitutivamente expresso, sendo empregado como gene controle para estudos de expressão gênica por RT-qPCR (Vandesompele et al., 2002). Em *Plasmodium falciparum* o gene de ubiquitina é constitutivamente expresso ao longo do ciclo de vida do parasita (Le Roch et al., 2003).

Para realização dos experimentos de RT-qPCR (item 3.19), 13 genes diferencialmente expressos foram selecionados: gene codificante para proteína de choque térmico de 90 kDa, proteína de micronema 2, serpina, dinamina, aspartil proteinase, aminopeptidase, proteína rombóide, transhidrogenase, BT1, antígeno não transmembranar, proteína de desenvolvimento do esporozoítos, 1 proteína putativa de *E. tenella* e 1 proteína hipotéticas. Em *E. tenella*, *E. acervulina* e/ou *E. bovis*, a maioria destas proteínas tem grande importância na invasão e manutenção do parasita, sendo essenciais para o esporozoítos, estágio formado ao final da esporulação. Pouco se conhece sobre a importância destas proteínas em *E. maxima*.

As proteínas de choque térmico (*Heat shock protein* - HSP) são chaperonas moleculares importantes para a homeostasia e proteção das células aos efeitos nocivos do estresse ambiental, sendo observada em maior concentração em situações de estresse celular e/ou oxidativo (Maresca e Carratu, 1992). Em parasitas, esta proteína tem importante função, já que durante a invasão, estabelecimento e multiplicação na célula hospedeira, estes organismos são expostos a várias condições de estresse (Newport et al., 1988; Maresca e Carratu, 1992). Particularmente a HSP90, proteína de choque térmico com tamanho

aproximado de 90 kDa (Péroval et al., 2006), além de ter importante papel na resposta ao estresse, também está envolvida com a regulação do desenvolvimento do parasita (Echeverria et al., 2005). As principais funções da HSP90 estão relacionadas à sinalização celular e manutenção da conformação adequada das proteínas, auxiliando no dobramento das cadeias polipeptídicas nascentes e redobramento das proteínas desnaturadas após o estresse (Nathan et al., 1997). Em *Toxoplasma gondii* (Echeverria et al., 2005) e *Eimeria tenella*, esta proteína tem grande importância no processo de invasão e multiplicação do parasita dentro da célula hospedeira (Péroval et al., 2006). Em *E. tenella*, esta chaperona também é importante para a manutenção da homeostase das fases intracelulares do parasita (Miska et al., 2005).

Sabe-se que a HSP90 é expressa em distintas fases do ciclo biológico de *Eimeria* spp., incluindo oocistos em diferentes fases da esporulação (Clark et al., 1996; Miska et al., 2005). Em relação aos nossos resultados de expressão gênica, foi observado tanto por ORESTES quanto por RT-qPCR um perfil diferenciado entre as três fases da esporulação estudadas, ocorrendo uma maior transcrição deste gene em oocistos esporulados. Estudos sobre a transcrição de HSP90 em demais espécies de *Eimeria* mostraram que em *E. bovis*, a transcrição deste gene foi detectada em merozoítos e em oocistos esporulados (Clark et al., 1996). Em *E. tenella* a HSP90 é aparentemente constitutiva, pois é transcrita em todas as fases do ciclo biológico, inclusive oocistos em diferentes fases da esporogonia, apresentando grande nível de transcrição após 48 horas do início da esporulação. Já para *E. acervulina*, este gene é altamente transcrito em oocistos não esporulados, não sendo detectado após 3 horas do início da esporulação e é novamente observado em esporozoítos e merozoítos (Miska et al., 2005). Interessante notar que apesar da HSP90 de *E. maxima*, *E. acervulina* e *E. tenella* terem sequências similares, o perfil de expressão gênica durante a esporulação nestas três espécies é diferenciado, indicando que a regulação da transcrição deste gene pode ser diferencial entre as espécies.

As micronemas são organelas localizadas na porção apical das fases invasivas de todos os parasitas do Filo Apicomplexa (Dubremetz et al., 1998). Estas organelas tem papel fundamental na adesão e invasão das células hospedeiras (Ryan et al., 2000), além de serem importantes para a motilidade do parasita e reconhecimento da célula alvo, sendo a primeira organela a liberar seu conteúdo

durante o processo de invasão (Bromley et al., 2003). As micronemas também estão envolvidas no tráfico e sequestro de ligantes para os receptores da célula hospedeira, liberando altas concentrações de adesinas na extremidade apical do parasita após contato com a célula alvo (Tomley e Soldati, 2001). Para o Filo Apicomplexa, 29 proteínas de micronema foram identificadas (Tomley e Soldati, 2001) e a abundância destas varia de acordo com o gênero e a espécie do protozoário (Carruthers e Tomley, 2008). Até o momento, em *Eimeria*, foram descritas e caracterizadas sete diferentes tipos de proteínas de micronema (Tomley e Soldati, 2001), as quais apresentam diferentes motivos de adesinas, tais como: *I-domains* (integrinas), repetições de TSP-1 (trombospondina tipo 1), repetições de TSP degenerada, repetições de *EGF-like* e *Apple domains* (Tomley e Soldati, 2001).

Não há muitas informações sobre a proteína de micronema 2, sabe-se que tem um tamanho de 50 kDa (Tomley et al., 1996) e o gene codificante está localizado no cromossomo 9 (Ryan et al., 2000). É relatado que esta proteína é secretada para a superfície do esporozoíto, porém sua função específica ainda não é conhecida, acredita-se que tenha grande importância na interação parasita hospedeiro como as demais proteínas de micronema (Tomley et al., 1996).

Ao analisar a transcrição do gene codificante para proteína de micronema 2 nas diferentes fases da esporulação, observamos que para ambas as metodologias testadas (ORESTES e RT-qPCR) a maior transcrição ocorreu em oocistos esporulados, seguida por parcialmente esporulados e não esporulados, indicando uma tendência de aumento da transcrição ao longo da esporulação. Estudos de transcrição do gene de micronema 2 em *E. tenella* mostraram que o gene codificante para esta proteína é transcrito em oocistos esporulados, não sendo detectado em oocistos não esporulados (Miska et al., 2004). Ryan et al. (2000) avaliaram a transcrição dos genes codificantes para as micronemas 1 a 5 ao longo da esporulação em *E. tenella* e observaram que a partir das 12 horas do início da esporulação, o gene codificante para micronema 2 começa a ser transcrito. Até o momento, não há dados na literatura acerca da transcrição deste gene nas diferentes fases evolutivas do ciclo de *E. maxima*. Comparando-se os dados de transcrição em oocistos de *E. maxima* com os dados de *E. tenella*, notamos que para micronema 2, o perfil de expressão gênica durante a esporulação parece ser semelhante entre ambas as espécies. O fato do gene da micronema 2 ser transcrito ao longo da esporulação com maiores níveis em oocistos esporulados talvez esteja

relacionado ao processo de formação dos esporozoítos, já que são proteínas importantes para a invasão do parasita.

Proteínas rombóides (ROM) são proteases intramembranares da classe das serinas (Urban et al., 2001). Os genes codificantes para estas proteínas estão presentes no genoma de muitos parasitas do Filo Apicomplexa, como por exemplo, *Toxoplasma gondii*, *Plasmodium* spp., *Eimeria tenella*, *Cryptosporidium* spp. e *Theileria* spp (Dowse e Soldati, 2005). Durante a invasão da célula alvo, estas proteínas estão envolvidas com a sinalização celular e clivagem de proteínas de micronema (Dowse et al., 2005; Freeman, 2008). Além de serem essenciais para a invasão, também participam do processo de internalização do parasita na célula hospedeira e formação do vacúolo parasitóforo (Freeman, 2008).

Ao analisar os resultados de expressão do gene codificante para a proteína rombóide em oocistos de *E. maxima* observamos que tanto para ORESTES quanto para RT-qPCR, este gene é transcrito em oocistos não esporulados, parcialmente esporulados e esporulados. De modo semelhante, em *E. tenella*, foi verificado por RT-PCR que há transcrição deste gene durante toda as fases da esporulação estudadas (0, 12, 24, 48 e 72 horas de esporulação) (Fetterer et al., 2007b). Em *E. maxima*, transcritos referentes a estes gene foram encontrados em oocistos esporulados, não sendo avaliada a presença destes transcritos nas demais fase da esporulação (Dong et al., *in press*).

Analisando o perfil de expressão do gene codificante para a proteína rombóide nas três fases da esporulação de *E. maxima* nota-se por RT-qPCR, que este gene é mais transcrito em oocistos parcialmente esporulados, seguido por não esporulados e esporulados.

Em estudo do transcriptoma de oocistos de *E. tenella* empregando hibridização subtrativa foi observado que a proteína rombóide é transcrita tanto em oocistos esporulados quanto em oocistos não esporulados, dados de RT-qPCR mostraram que a transcrição deste gene é maior em oocistos não esporulados do que em oocistos esporulados (Miska et al., 2004). Como este trabalho não avaliou a transcrição deste gene na fase de esporoblasto, não foi possível fazer uma análise comparativa dos perfis de expressão nas diferentes fases da esporulação entre *E. maxima* e *E. tenella*. Se analisarmos somente a transcrição em oocistos não esporulados e esporulados verificamos um perfil semelhante entre estas espécies de *Eimeria*.

Serpinas é uma superfamília de proteínas cuja função de um modo geral, é de inibir serino proteases e por regularem cascatas proteolíticas, são essenciais para muitos processos fisiológicos (Huber e Carrell, 1989; Irving et al., 2000). Sugere-se que as serpinas tenham a função de modular a resposta inflamatória do hospedeiro, a secreção de proteinases pelos parasitas e proteger os parasitas quando são ingeridos pelo hospedeiro, já que o trato digestivo é rico em proteinases (Bruno et al., 2004).

Serpinas já foram descritas em organismos do Filo Apicomplexa (Pszenny et al., 2000; Pszenny et al., 2002; Bruno et al., 2004; Fetterer et al., 2008; Dong et al., *in press*). Em *T. gondii*, esta proteína tem grande importância para a sobrevivência de taquizoítos e invasão das células hospedeiras (Conseil et al., 1999). Em *Eimeria*, apesar da função desta proteína ainda não ser conhecida, acredita-se que a serpina seja secretada pelo parasita durante a invasão e em outros processos básicos do desenvolvimento do parasita (Fetterer et al., 2008).

O gene codificante para serpina foi identificado e caracterizado em esporozoítos de *E. acervulina* (Fetterer et al., 2008). Em *E. maxima*, transcritos referentes a estes genes foram encontrados em oocistos esporulados, não sendo avaliada a presença destes nas demais fases da esporulação (Dong et al., *in press*).

Após a análise comparativa dos dados de *E. maxima* com os obtidos pelo projeto ORESTES/ESTs para *E. tenella* e *E. acervulina*, notou-se que para as três espécies o gene codificador para serpina é mais transcrito em oocistos esporulados.

Segundo os resultados de RT-qPCR de *E. maxima* há uma maior transcrição do gene codificante para serpina em oocistos esporulados do que parcialmente esporulados e não esporulados, mostrando uma tendência de aumento da transcrição durante a esporulação. Resultados semelhantes foram descritos para *E. tenella*, onde a transcrição deste gene foi maior em oocistos esporulados do que não esporulados (Miska et al., 2004). Infelizmente, como neste estudo a transcrição deste gene não foi avaliada em oocistos parcialmente esporulados de *E. tenella*, não foi possível verificar se estas espécies teriam mesmo perfil de expressão.

Como o gene codificante de serpina foi observado em maiores níveis em oocistos esporulados, talvez a presença deste possa estar relacionada ao fato de que, assim que os esporozoítos existam no lúmen do intestino, deparam-se com um ambiente repleto de proteases, necessitando de inibidores destas para manterem-se viáveis.

A aminopeptidase é uma enzima da classe das metaloproteases que em parasitas do filo Apicomplexa tem significativa importância na maturação de proteínas, invasão celular e resposta imune (Okhuysen et al., 1996; Berthonneau et al., 2000; Padda et al., 2002) De fato, foi observado que inibidores de aminopeptidase reduzem a invasão celular e a multiplicação *in vitro* de taquizoítos de *Toxoplasma gondii* (Conseil et al., 1999) e também bloqueiam a excitação de esporozoítos de *Cryptosporidium parvum* (Okhuysen et al., 1996).

Nossos dados revelaram uma maior transcrição do gene codificante para aminopeptidase em oocistos não esporulados, seguido de oocisto parcialmente esporulado e esporulado, indicando que a transcrição deste gene diminui ao longo da esporulação. Estudos anteriores mostraram que a aminopeptidase também é transcrita em oocistos de *E. tenella* (Miska et al., 2004; Fetterer et al., 2005). Apesar de Miska et al. (2004) terem observado por hibridização subtrativa que a aminopeptidase é transcrita somente em oocistos esporulados, em trabalho subsequente sobre a caracterização de aminopeptidase em *E. tenella*, a transcrição deste gene foi detectada ao longo de toda a esporulação tendendo a decair à medida que os esporozoítos eram formados (Fetterer et al., 2005). Nossos resultados são semelhantes a este trabalho, indicando que o perfil de transcrição deste gene durante a esporulação deve ser semelhante entre *E. maxima* e *E. tenella*. Especula-se que esta enzima deve estar envolvida com o processamento de proteínas durante a esporulação (Fetterer et al., 2005).

Aspartil proteinases são enzimas que tem conformação bilobal e dois resíduos de aspartato no sítio catalítico (Tang e Wong, 1987; Jean et al., 2001b). Esta proteína é altamente conservada entre as diversas espécies (Jean et al., 2001a) e está relacionada à patogenicidade e sobrevivência de vários organismos, incluindo vírus, fungos e protozoários (Jean et al., 2000). Em *Eimeria tenella*, esta enzima é conhecida como eimepsina, localiza-se no corpo refrátil dos esporozoítos e é codificada por um gene de cópia única (Jean et al., 2000).

Os nossos dados mostram, para ambas as técnicas empregadas (ORESTES e RT-qPCR), que a transcrição do gene codificante para eimepsina é observada nas três fases estudadas, sendo maior em oocistos parcialmente esporulados, seguido de oocistos não esporulados e esporulados. Miska et al. (2004) relataram que em *E. tenella* este gene é transcrito tanto em oocistos não esporulados quanto em oocistos esporulados, não avaliando a transcrição em oocistos parcialmente esporulados. De

modo semelhante ao observado para *E. maxima*, este gene é transcrito ao longo da esporulação em *E. tenella*, onde a maior expressão foi observada na fase de esporoblasto tendendo a declinar após esta fase (Jean et al., 2001b).

É descrito que em *E. tenella* a eimepsina está presente nos corpos refrateis dos esporozoítos. Estudos de imuno-localização proteica mostraram que a eimepsina passa a ser detectada no complexo apical quando os esporozoítos entram em contato com as células do hospedeiro, sugerindo que esta enzima é importante para o processo invasivo (Jean et al., 2000).

A dinamina é uma GTPase responsável por diversos processos celulares, incluindo citocinese, divisão de organelas, formação de vesículas de transporte e resistência parasitária (Praefcke e McMahon, 2004; Breinich et al., 2009). Em parasitas do Filo Apicomplexa, esta proteína está relacionada à invasão da célula hospedeira, uma vez que estão envolvidas na formação de micronemas e roptrias (Li et al., 2004a; Charneau et al., 2007; Breinich et al., 2009; Webster e McFadden, 2009). Em *Toxoplasma gondii*, a dinamina é detectada em taquizoítos durante a endodiogenia, momento no qual estas organelas são formadas (Breinich et al., 2009). Em *Plasmodium falciparum* o gene codificante de dinamina é transcrito durante a morfogênese dos merozoítos (Florent et al., 2004). Embora tenha papel importante na invasão e sobrevivência de parasitas do Filo Apicomplexa, não há, até o momento, nenhuma descrição ou caracterização do gene codificante para dinamina em *Eimeria spp.* Experimentos de RT-qPCR mostram que este gene é transcrito em todas as fases estudadas, apresentando maior expressão em oocistos não esporulados. Analisando as três fases estudadas, nota-se que a transcrição deste gene tende a diminuir ao longo da esporulação.

Baseando-se nos relatos para *Toxoplasma gondii* (Breinich et al., 2009), a transcrição da dinamina durante a esporulação poderia estar relacionada à gênese de micronemas e roptrias durante a formação dos esporozoítos.

O BT1 (transportador de biopterina) é um transportador de folatos, os quais estão envolvidos na produção de purinas e pirimidinas para a replicação de DNA, bem como, síntese e catabolismo de vários aminoácidos, tais como: metionina, glicina, serina, glutamato e histidina (Massimine et al., 2005). Embora os parasitas do Filo Apicomplexa tenham a capacidade de sintetizar folato *de novo*, mecanismos de transporte de folatos também têm sido descritos nestes protozoários. Como estes componentes são essenciais para o crescimento do parasita, talvez o transporte de

folatos seja utilizado nas diferentes fases do ciclo de vida; sob diferentes condições ambientais, onde os folatos exógenos são particularmente abundantes e nos casos em que a biossíntese de folato *de novo* não é suficiente para a sobrevivência do parasita (Jansen et al., 1990; Said et al., 2000; Massimine et al., 2005)

O gene BT1 já foi estudado em *Leishmania donovani* (Dole et al., 2002) e em *Toxoplasma gondii* (Massimine et al., 2005). Até o momento não há estudos que avaliem a transcrição, identificação e caracterização do gene codificante para proteínas da família BT1 em *Eimeria* spp. Os perfis de expressão gênica obtido por ORESTES/ESTs foram bastante similares para as três espécies, o gene é transcrito em todas as fases estudadas e há uma tendência de queda da transcrição durante a esporulação. Ao avaliar os resultados de expressão gênica por RT-qPCR observa-se uma maior transcrição deste gene em oocistos parcialmente esporulados. Transcrição do gene codificador de BT1 foi observado somente em oocistos não esporulados de *E. tenella*, não sendo detectado em oocistos esporulados (Miska et al., 2004).

Considerando que os folatos são essenciais para a manutenção e sobrevivência de vários organismos, o estudo destes transportadores poderia contribuir para um melhor conhecimento da via metabólica do ácido fólico. Inibidores da síntese de folatos, como pirimetamina (Mathis et al., 1984) e sulfonamida (Campbell, 2008) são importantes quimioterápicos usados para controle da coccidiose aviária.

A transhidrogenase é uma enzima responsável pela translocação de íons de hidrogênio entre NAD(H) e NADP(H) na membrana de mitocôndria e de bactérias (Jackson, 2003). Em protozoários, a sequência do gene codificante para transhidrogenase já foi descrita em *Plasmodium falciparum*, *Entamoeba histolytica* e *E. tenella* (Kramer et al., 1993). Há poucos trabalhos reportando a presença desta enzima em parasitas do gênero *Eimeria* e acredita-se que nestes parasitas, a transidrogenase pode estar relacionada a hidrólise de ATP e respiração celular (Vermeulen et al., 1993). Em *E. maxima*, transcritos referentes a estes gene foram encontrados em oocistos esporulados, não sendo avaliada a presença destes nas demais fases da esporulação (Dong et al., *in press*).

Avaliando a transcrição do gene codificante para a transhidrogenase em oocistos de *E. maxima* observamos que em ambas as técnicas estudadas (ORESTES e RT-qPCR) este gene é transcrito nas três fases da esporulação, sendo

mais expresso em oocistos parcialmente esporulados, seguidos de oocistos não esporulados e esporulados. Em *E. tenella*, este gene é transcrito em oocistos não esporulados e esporulados (Miska et al., 2004), porém não há dados referentes à transcrição deste gene em oocistos parcialmente esporulados. Nossos dados de ORESTES/ESTs mostram que o perfil de expressão em *E. tenella* é bastante similar ao de *E. maxima*.

Apesar de não existir dados, até o momento, sobre a expressão gênica e função relacionadas à proteína de desenvolvimento de esporozoíto, antígeno transmembranar, proteína putativa não caracterizada de *E. tenella* e proteína hipotética, os transcritos referentes a estes produtos foram escolhidos para estudo de RT-qPCR pois estão presentes no grupo dos 100 transcritos diferencialmente expressos mais abundantes.

Analisando os perfis de expressão de ORESTES/ESTs para os genes codificantes de proteína de desenvolvimento do esporozoíto e para proteína putativa de *Eimeria tenella* observa-se para as três espécies de *Eimeria*, a presença de transcritos destes genes nas três fases da esporulação. Ao analisar o perfil de expressão gênica em *E. maxima*, observamos que para o gene codificante da proteína putativa, o perfil de expressão é similar ao obtido para *E. tenella*, enquanto que para o gene codificante para proteína de desenvolvimento do esporozoíto, o perfil de expressão é semelhante ao de *E. acervulina*. Em *E. maxima*, para ambos os genes, os resultados de RT-qPCR, mostraram que a transcrição em oocistos parcialmente esporulados foi maior do que a mensurada em oocistos não esporulados e esporulados.

Em relação à proteína de desenvolvimento de esporozoíto, os resultados de Interpro e de RPS-BLAST mostraram a presença de domínio de insulínase, peptidase da família M16. A insulínase é uma metaloproteinase que além de clivar insulina, está envolvida com vários processos celulares (Fernandez-Gamba et al., 2009). Abrahamsen et al. (1993) também detectaram a presença de transcrito similar em oocistos esporulados de *Eimeria bovis*. Em oocistos, talvez esta peptidase esteja envolvida com a degradação de proteínas relacionadas à regulação da esporulação.

O resultado de Interpro para a proteína não caracterizada de *Eimeria tenella* mostra a presença de sítio conservado de EGF (Epidermal Growth Factor). Proteínas contendo repetições de EGF já foram descritas em oocistos de *Toxoplasma gondii*, *Cryptosporidium parvum* e em cisto de *Giardia lamblia* (Chiu et

al., 2010). Ao que parece, estes produtos devem estar relacionados à proteína de parede de oocisto (Possenti et al., 2010).

Comparando os perfis de ORESTES de *E. maxima* com os de *E. acervulina*, para os genes codificantes para antígeno não transmembranar e proteína hipotética, observamos que para ambos os genes e espécies, o estágio de maior transcrição é na fase de oocistos esporulados. Ao analisar os resultados obtidos por RT-qPCR, os dois genes são mais transcritos em oocistos esporulados, seguidos de oocistos parcialmente esporulados e não esporulados, mostrando uma tendência de declínio da transcrição durante a esporulação.

Analisando os resultados de Interpro e RPS-BLAST para o antígeno transmembranar observamos que este gene está relacionado ao metabolismo e/ou transporte de carboidratos. Como já foi descrito transporte de glicose em esporozoítos de *Eimeria tenella* (Smith e Lee, 1986) e considerando que a esporulação culmina na formação do esporozoíto, provavelmente este gene esteja relacionado ao transporte destes açúcares.

Em relação à proteína hipotética, foi observada tanto para ORESTES quanto para RT-qPCR, uma transcrição preferencial em oocistos esporulados. Para RT-qPCR a transcrição nestes oocistos é 168 vezes maior do que oocistos não esporulados. Após análise no programa Threader, observamos que esta proteína apresentou enovelamento predito similar a treonina aspartase 1. Até o momento, pouco se sabe sobre esta endopeptidase. Em humanos, é denominada taspase1 a qual pertence à uma classe distinta de enzimas proteolíticas, sendo responsável por clivar vários substratos intracelulares (Hsieh et al., 2003).

Ao comparar os resultados de expressão gênica diferencial de *Eimeria maxima* obtidos por ORESTES e qRT-PCR, observamos que os perfis de expressão foram semelhantes para 62% dos transcritos, indicando uma boa concordância entre as técnicas. Para os diferentes estágios, observamos de um modo geral que os resultados de transcrição para oocistos esporulados foram mais consistentes e concordantes entre as duas técnicas do que os de oocistos parcialmente esporulados e não esporulados. Como foi gerado um número menor de leituras para estes dois últimos estágios, acreditamos que a geração de um número maior de sequências destas bibliotecas poderia minimizar estas diferenças.

A utilização da metodologia de ORESTES em organismos do gênero *Eimeria* é inédita e permitiu a análise da expressão diferencial dos transcritos provenientes

de três estágios da esporulação de oocistos de *Eimeria maxima*. Além de disso, viabilizou a escolha de candidatos interessantes para estudos futuros relacionados à expressão gênica durante a esporulação.

O estudo da expressão gênica em oocistos poderá no futuro corroborar para um melhor conhecimento dos genes envolvidos com a viabilidade dos oocistos no ambiente, formação dos esporozoítos bem como sua excitação, passo este fundamental para o estabelecimento da infecção.

6 CONCLUSÃO

- Foram construídas bibliotecas ORESTES, a partir de oocistos não esporulados, parcialmente esporulados e esporulados, o que resultou no sequenciamento de 20.149 leituras, as quais, após a montagem, geraram 1.207 *contigs* e 2.091 *singlets*.
- Menos de 2% dos *contigs* encontrados são de alta expressão (>100x), o que significa que um pequeno conjunto de genes é altamente expresso em cada fase do parasita.
- A maioria (73,5%) das leituras dos *contigs* era proveniente de uma única fase da esporulação, indicando assim que a transcrição gênica é estágio-específica.
- Os resultados obtidos pelo agrupamento dos perfis gênicos em função dos estágios mostraram que oocistos agrupam num único clado e merozoítos num clado externo, sendo condizentes com a biologia do parasita.
- Após a análise estatística, foram obtidos 379 *contigs* diferencialmente expressos e 828 não diferencialmente expressos.
- Menos de 30% dos transcritos anotados apresentou resultados positivos de BLAST, o que pode indicar que um grande número de genes de *E.maxima* ainda está por ser caracterizado.
- O protocolo de anotação automática gerou dados consistentes, 74,2% dos transcritos com resultados de BLAST positivos apresentaram similaridade com sequências de organismos do Filo Apicomplexa.
- Após análise no programa Threader 3, aproximadamente 72% das proteínas hipotéticas presentes entre os 100 transcritos diferencialmente expressos mais abundantes, apresentaram evovelamento similar com proteínas depositadas no banco de dados RCSB PDB.

- Dentre os 100 transcritos diferencialmente expressos mais abundantes, 10 apresentaram perfil de transcrição gênica em oocistos semelhante à de *E. acervulina* e outras 13 à de *E. tenella*.
- Um pequeno conjunto de genes diferencialmente expressos entre as três fases da esporulação foi submetido a experimentos de RT-qPCR, sendo observada boa concordância entre os perfis de expressão gênica obtidos por ORESTES e PCR em tempo real.

REFERÊNCIAS

REFERÊNCIAS*

ABEF (Associação Brasileira dos Produtores e Exportadores de Frangos). Relatório Anual 2009/2010. 2009. Available from: http://www.abef.com.br/portal/_clientes/abef/cat/RA_2010.pdf [2010 dez. 10].

Abrahamsen MS, Clark TG, Mascolo P, Speer CA, White MW. Developmental gene expression in *Eimeria bovis*. *Mol Biochem Parasitol*. 1993 Jan;57(1):1-14.

Abrahamsen MS, Johnson RR, Clark TG, White MW. Developmental regulation of an *Eimeria bovis* mRNA encoding refractile body-associated proteins. *Mol Biochem Parasitol*. 1994 Nov;68(1):25-34.

Abrahamsen MS, Johnson RR, Hathaway M, White MW. Identification of *Eimeria bovis* merozoite cDNAs using differential mRNA display. *Mol Biochem Parasitol*. 1995 May;71(2):183-91.

Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, et al. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science*. 1991 Jun 21;252(5013):1651-6.

Allen PC, Fetterer RH. Recent advances in biology and immunobiology of *Eimeria* species and in diagnosis and control of infection with these coccidian parasites of poultry. *Clin Microbiol Rev*. 2002 Jan;15(1):58-65.

Allen PC, Jenkins MC, Miska KB. Cross protection studies with *Eimeria maxima* strains. *Parasitol Res*. 2005 Oct;97(3):179-85.

Allocco JJ, Profous-Juchelka H, Myers RW, Nare B, Schmatz DM. Biosynthesis and catabolism of mannitol is developmentally regulated in the protozoan parasite *Eimeria tenella*. *J Parasitol*. 1999 Apr;85(2):167-73.

Almeida CR, Stoco PH, Wagner G, Sincero TC, Rotava G, Bayer-Santos E, Rodrigues JB, Sperandio MM, Maia AA, Ojopi EP, Zaha A, Ferreira HB, Tyler KM, Davila AM, Grisard EC, Dias-Neto E. Transcriptome analysis of *Taenia solium*

*De acordo com:

International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to Biomedical Journal: Sample references. Available from: <http://www.icmje.org> [2007 May 22].

cysticerci using Open Reading Frame ESTs (ORESTES). *Parasit Vectors*. 2009;2(1):35.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997 Sep 1;25(17):3389-402.

Arnaldi LA, Borra RC, Maciel RM, Cerutti JM. Gene expression profiles reveal that DCN, DIO1, and DIO2 are underexpressed in benign and malignant thyroid tumors. *Thyroid*. 2005 Mar;15(3):210-21.

Audic S, Claverie JM. The significance of digital gene expression profiles. *Genome Res*. 1997 Oct;7(10):986-95.

Aurrecochea C, Heiges M, Wang H, Wang Z, Fischer S, Rhodes P, Miller J, Kraemer E, Stoeckert CJ, Roos DS, Kissinger JC. ApiDB: integrated resources for the apicomplexan bioinformatics resource center. *Nucleic Acids Research*. 2007 Jan;35:D427-D30.

Avedissian M, Longo BM, Jaqueta CB, Schnabel B, Paiva PB, Mello LE, Briones MR. Hippocampal gene expression analysis using the ORESTES methodology shows that homer 1a mRNA is upregulated in the acute period of the pilocarpine epilepsy model. *Hippocampus*. 2007;17(2):130-6.

Barta JR, Coles BA, Schito ML, Fernando MA, Martin A, Danforth HD. Analysis of infraspecific variation among five strains of *Eimeria maxima* from North America. *International Journal for Parasitology*. 1998 Mar;28(3):485-92.

Basak SC, Lee S, Barta JR, Fernando MA. Differential display analysis of gene expression in two immunologically distinct strains of *Eimeria maxima*. *Parasitol Res*. 2006 Jun;99(1):28-36.

Belli SI, Ferguson DJ, Katrib M, Slapetova I, Mai K, Slapeta J, Flowers SA, Miska KB, Tomley FM, Shirley MW, Wallach MG, Smith NC. Conservation of proteins involved in oocyst wall formation in *Eimeria maxima*, *Eimeria tenella* and *Eimeria acervulina*. *Int J Parasitol*. 2009 Aug;39(10):1063-70.

Belli SI, Walker RA, Flowers SA. Global protein expression analysis in apicomplexan parasites: current status. *Proteomics*. 2005 Mar;5(4):918-24.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol*. 2004 Jul 16;340(4):783-95.

Berthonneau J, Rodier MH, El Moudni B, Jacquemin JL. *Toxoplasma gondii*: purification and characterization of an immunogenic metallopeptidase. *Exp Parasitol.* 2000 Jun;95(2):158-62.

Bishop R, Shah T, Pelle R, Hoyle D, Pearson T, Haines L, Brass A, Hulme H, Graham SP, Taracha EL, Kanga S, Lu C, Hass B, Wortman J, White O, Gardner MJ, Nene V, de Villiers EP. Analysis of the transcriptome of the protozoan *Theileria parva* using MPSS reveals that the majority of genes are transcriptionally active in the schizont stage. *Nucleic Acids Res.* 2005;33(17):5503-11.

Blake DP, Hesketh P, Archer A, Carroll F, Shirley MW, Smith AL. The influence of immunizing dose size and schedule on immunity to subsequent challenge with antigenically distinct strains of *Eimeria maxima*. *Avian Pathol.* 2005 Dec;34(6):489-94.

Blake DP, Oakes R, Smith AL. A genetic linkage map for the apicomplexan protozoan parasite *Eimeria maxima* and comparison with *Eimeria tenella*. *Int J Parasitol.* 2011 Oct 7;41(2):263-70.

Blake DP, Qin Z, Cai J, Smith AL. Development and validation of real-time polymerase chain reaction assays specific to four species of *Eimeria*. *Avian Pathol.* 2008 Feb;37(1):89-94.

Boguski MS, Lowe TMJ, Tolstoshev CM. Dbest - Database for Expressed Sequence Tags. *Nature Genetics.* 1993 Aug;4(4):332-3.

Bozdech Z, Zhu J, Joachimiak MP, Cohen FE, Pulliam B, DeRisi JL. Expression profiling of the schizont and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray. *Genome Biol.* 2003;4(2):R9.

Brackett S, Bliznick A. The occurrence and economic importance of coccidiosis in chickens. New York: American Cyanamid Co.; 1950.

Brady SM, Long TA, Benfey PN. Unraveling the dynamic transcriptome. *Plant Cell.* 2006 Sep;18(9):2101-11.

Breinich MS, Ferguson DJ, Foth BJ, van Dooren GG, Lebrun M, Quon DV, Striepen B, Bradley PJ, Frischknecht F, Carruthers VB, Meissner M. A dynamin is required for the biogenesis of secretory organelles in *Toxoplasma gondii*. *Curr Biol.* 2009 Feb 24;19(4):277-86.

Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M, Roth R, George D, Eletr S, Albrecht G, Vermaas E, Williams SR, Moon K, Burcham T, Pallas M, DuBridgbe RB, Kirchner J, Fearon K, Mao J, Corcoran K. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat Biotechnol.* 2000 Jun;18(6):630-4.

Brentani H, Caballero OL, Camargo AA, da Silva AM, da Silva WA, Jr., Dias Neto E, Grivet M, Gruber A, Guimaraes PE, Hide W, Iseli C, Jongeneel CV, Kelso J, Nagai MA, Ojopi EP, Osorio EC, Reis EM, Riggins GJ, Simpson AJ, de Souza S, Stevenson BJ, Strausberg RL, Tajara EH, Verjovski-Almeida S, Acencio ML, Bengtson MH, Bettoni F, Bodmer WF, Briones MR, Camargo LP, Cavenee W, Cerutti JM, Coelho Andrade LE, Costa dos Santos PC, Ramos Costa MC, da Silva IT, Estecio MR, Sa Ferreira K, Furnari FB, Faria M, Jr., Galante PA, Guimaraes GS, Holanda AJ, Kimura ET, Leerkes MR, Lu X, Maciel RM, Martins EA, Massirer KB, Melo AS, Mestriner CA, Miracca EC, Miranda LL, Nobrega FG, Oliveira PS, Paquola AC, Pandolfi JR, Campos Pardini MI, Passetti F, Quackenbush J, Schnabel B, Sogayar MC, Souza JE, Valentini SR, Zaiats AC, Amaral EJ, Arnaldi LA, de Araujo AG, de Bessa SA, Bicknell DC, Ribeiro de Camaro ME, Carraro DM, Carrer H, Carvalho AF, Colin C, Costa F, Curcio C, Guerreiro da Silva ID, Pereira da Silva N, Dellamano M, El-Dorry H, Espreadico EM, Scattoni Ferreira AJ, Ayres Ferreira C, Fortes MA, Gama AH, Giannella-Neto D, Giannella ML, Giorgi RR, Goldman GH, Goldman MH, Hackel C, Ho PL, Kimura EM, Kowalski LP, Krieger JE, Leite LC, Lopes A, Luna AM, Mackay A, Mari SK, Marques AA, Martins WK, Montagnini A, Mourao Neto M, Nascimento AL, Neville AM, Nobrega MP, O'Hare MJ, Otsuka AY, Ruas de Melo AI, Paco-Larson ML, Guimaraes Pereira G, Pesquero JB, Pessoa JG, Rahal P, Rainho CA, Rodrigues V, Rogatto SR, Romano CM, Romeiro JG, Rossi BM, Rusticci M, Guerra de Sa R, Sant' Anna SC, Sarmazo ML, Silva TC, Soares FA, Sonati Mde F, de Freitas Sousa J, Queiroz D, Valente V, Vettore AL, Villanova FE, Zago MA, Zalcberg H. The generation and utilization of a cancer-oriented representation of the human transcriptome by using expressed sequence tags. *Proc Natl Acad Sci U S A.* 2003 Nov 11;100(23):13418-23.

Bromley E, Leeds N, Clark J, McGregor E, Ward M, Dunn MJ, Tomley F. Defining the protein repertoire of microneme secretory organelles in the apicomplexan parasite *Eimeria tenella*. *Proteomics.* 2003 Aug;3(8):1553-61.

Bruno S, Duschak VG, Ledesma B, Ferella M, Andersson B, Guarnera EA, Angel SO. Identification and characterization of serine proteinase inhibitors from *Neospora caninum*. *Mol Biochem Parasitol.* 2004 Jul;136(1):101-7.

Bumstead J, Tomley F. Induction of secretion and surface capping of microneme proteins in *Eimeria tenella*. *Mol Biochem Parasitol.* 2000 Oct;110(2):311-21.

Cai X, Fuller AL, McDougald LR, Zhu G. Apicoplast genome of the coccidian *Eimeria tenella*. *Gene.* 2003 Dec 4;321:39-46.

Camargo AA, Samaia HP, Dias-Neto E, Simao DF, Migotto IA, Briones MR, Costa FF, Nagai MA, Verjovski-Almeida S, Zago MA, Andrade LE, Carrer H, El-Dorry HF, Espreafico EM, Habr-Gama A, Giannella-Neto D, Goldman GH, Gruber A, Hackel C, Kimura ET, Maciel RM, Marie SK, Martins EA, Nobrega MP, Paco-Larson ML, Pardini MI, Pereira GG, Pesquero JB, Rodrigues V, Rogatto SR, da Silva ID, Sogayar MC, Sonati MF, Tajara EH, Valentini SR, Alberto FL, Amaral ME, Aneas I, Arnaldi LA, de Assis AM, Bengtson MH, Bergamo NA, Bombonato V, de Camargo ME, Canevari RA, Carraro DM, Cerutti JM, Correa ML, Correa RF, Costa MC, Curcio C, Hokama PO, Ferreira AJ, Furuzawa GK, Gushiken T, Ho PL, Kimura E, Krieger JE, Leite LC, Majumder P, Marins M, Marques ER, Melo AS, Melo MB, Mestriner CA, Miracca EC, Miranda DC, Nascimento AL, Nobrega FG, Ojopi EP, Pandolfi JR, Pessoa LG, Prevedel AC, Rahal P, Rainho CA, Reis EM, Ribeiro ML, da Ros N, de Sa RG, Sales MM, Sant'anna SC, dos Santos ML, da Silva AM, da Silva NP, Silva WA, Jr., da Silveira RA, Sousa JF, Stecconi D, Tsukumo F, Valente V, Soares F, Moreira ES, Nunes DN, Correa RG, Zalcborg H, Carvalho AF, Reis LF, Brentani RR, Simpson AJ, de Souza SJ. The contribution of 700,000 ORF sequence tags to the definition of the human transcriptome. *Proc Natl Acad Sci U S A*. 2001 Oct 9;98(21):12103-8.

Campbell WC. History of the discovery of sulfaquinoxaline as a coccidiostat. *J Parasitol*. 2008 Aug;94(4):934-45.

Canning EU, Anwar M. Studies on meiotic division in coccidial and malarial parasites. *J Protozool*. 1968 May;15(2):290-8.

Carruthers VB, Tomley F. Receptor-ligand interaction and invasion: Microneme proteins in apicomplexans. *Subcellular Biochemistry* 2008;47:33-45.

Castañón CAB. Análise e reconhecimento digital de formas biológicas para o diagnóstico automático de parasitas do gênero *Eimeria*. São Paulo: Universidade de São Paulo; 2006.

Castañón CAB, Fraga JS, Fernandez S, Gruber A, Costa LD. Biological shape characterization for automatic image recognition and diagnosis of protozoan parasites of the genus *Eimeria*. *Pattern Recognition*. 2007 Jul;40(7):1899-910.

Chapman HD. Origins of coccidiosis research in the fowl - The first fifty years. *Avian Diseases*. 2003 Jan-Mar;47(1):1-20.

Chapman HD, Cherry TE, Danforth HD, Richards G, Shirley MW, Williams RB. Sustainable coccidiosis control in poultry production: the role of live vaccines. *Int J Parasitol*. 2002 May;32(5):617-29.

Chapman HD, Shirley MW. The Houghton strain of *Eimeria tenella*: a review of the type strain selected for genome sequencing. *Avian Pathol.* 2003 Apr;32(2):115-27.

Charneau S, Bastos IM, Mouray E, Ribeiro BM, Santana JM, Grellier P, Florent I. Characterization of PfDYN2, a dynamin-like protein of *Plasmodium falciparum* expressed in schizonts. *Microbes Infect.* 2007 Jun;9(7):797-805.

Chen J, Sun M, Lee S, Zhou G, Rowley JD, Wang SM. Identifying novel transcripts and novel genes in the human genome by using novel SAGE tags. *Proc Natl Acad Sci U S A.* 2002 Sep 17;99(19):12257-62.

Chiu PW, Huang YC, Pan YJ, Wang CH, Sun CH. A novel family of cyst proteins with epidermal growth factor repeats in *Giardia lamblia*. *PLoS Negl Trop Dis.* 2010;4(5):e677.

Clark TG, Abrahamsen MS, White MW. Developmental expression of heat shock protein 90 in *Eimeria bovis*. *Mol Biochem Parasitol.* 1996 Jun;78(1-2):259-63.

Cloonan N, Grimmond SM. Transcriptome content and dynamics at single-nucleotide resolution. *Genome Biol.* 2008;9(9):234.

Conseil V, Soete M, Dubremetz JF. Serine protease inhibitors block invasion of host cells by *Toxoplasma gondii*. *Antimicrob Agents Chemother.* 1999 Jun;43(6):1358-61.

Current WL, Upton SJ, Long PL. Taxonomy and Life Cycles. In: Long PL, editor. *Coccidiosis of Man and Domestic Animals*. Boston: CRC Press Inc; 1990. p. 1-17.

da Mota AF, Sonstegard TS, Van Tassell CP, Shade LL, Matukumalli LK, Wood DL, Capuco AV, Brito MA, Connor EE, Martinez ML, Coutinho LL. Characterization of open reading frame-expressed sequence tags generated from *Bos indicus* and *B. taurus* mammary gland cDNA libraries. *Anim Genet.* 2004 Jun;35(3):213-9.

Dalloul RA, Lillehoj HS. Recent advances in immunomodulation and vaccination strategies against coccidiosis. *Avian Dis.* 2005 Mar;49(1):1-8.

de Souza SJ, Camargo AA, Briones MR, Costa FF, Nagai MA, Verjovski-Almeida S, Zago MA, Andrade LE, Carrer H, El-Dorry HF, Espreafico EM, Habr-Gama A, Giannella-Neto D, Goldman GH, Gruber A, Hackel C, Kimura ET, Maciel RM, Marie SK, Martins EA, Nobrega MP, Paco-Larson ML, Pardini MI, Pereira GG, Pesquero JB, Rodrigues V, Rogatto SR, da Silva ID, Sogayar MC, de Fatima Sonati M, Tajara

EH, Valentini SR, Acencio M, Alberto FL, Amaral ME, Aneas I, Bengtson MH, Carraro DM, Carvalho AF, Carvalho LH, Cerutti JM, Correa ML, Costa MC, Curcio C, Gushiken T, Ho PL, Kimura E, Leite LC, Maia G, Majumder P, Marins M, Matsukuma A, Melo AS, Mestriner CA, Miracca EC, Miranda DC, Nascimento AN, Nobrega FG, Ojopi EP, Pandolfi JR, Pessoa LG, Rahal P, Rainho CA, da Ros N, de Sa RG, Sales MM, da Silva NP, Silva TC, da Silva W, Jr., Simao DF, Sousa JF, Stecconi D, Tsukumo F, Valente V, Zalcbeg H, Brentani RR, Reis FL, Dias-Neto E, Simpson AJ. Identification of human chromosome 22 transcribed sequences with ORF expressed sequence tags. *Proc Natl Acad Sci U S A*. 2000 Nov 7;97(23):12690-3.

de Vries E, Corton C, Harris B, Cornelissen AW, Berriman M. Expressed sequence tag (EST) analysis of the erythrocytic stages of *Babesia bovis*. *Vet Parasitol*. 2006 May 31;138(1-2):61-74.

Deepak S, Kottapalli K, Rakwal R, Oros G, Rangappa K, Iwahashi H, Masuo Y, Agrawal G. Real-Time PCR: Revolutionizing Detection and Expression Analysis of Genes. *Curr Genomics*. 2007 Jun;8(4):234-51.

Del Cacho E, Gallego M, Pages M, Monteagudo L, Sanchez-Acedo C. HSP70 is part of the synaptonemal complex in *Eimeria tenella*. *Parasitol Int*. 2008 Dec;57(4):454-9.

Del Cacho E, Gallego M, Pereboom D, Lopez-Bernad F, Quilez J, Sanchez-Acedo C. *Eimeria tenella*: hsp70 expression during sporogony. *J Parasitol*. 2001 Oct;87(5):946-50.

Del Cacho E, Pages M, Gallego M, Barbero JL, Monteagudo L, Sanchez-Acedo C. Meiotic chromosome pairing and bouquet formation during *Eimeria tenella* sporulation. *Int J Parasitol*. 2010 Mar 15;40(4):453-62.

Del Cacho E, Pages M, Gallego M, Monteagudo L, Sanchez-Acedo C. Synaptonemal complex karyotype of *Eimeria tenella*. *Int J Parasitol*. 2005 Nov;35(13):1445-51.

Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques*. 2004 Jul;37(1):112-4, 6, 8-9.

Dias Neto E, Correa RG, Verjovski-Almeida S, Briones MR, Nagai MA, da Silva W, Jr., Zago MA, Bordin S, Costa FF, Goldman GH, Carvalho AF, Matsukuma A, Baia GS, Simpson DH, Brunstein A, de Oliveira PS, Bucher P, Jongeneel CV, O'Hare MJ, Soares F, Brentani RR, Reis LF, de Souza SJ, Simpson AJ. Shotgun sequencing of the human transcriptome with ORF expressed sequence tags. *Proc Natl Acad Sci U S A*. 2000 Mar 28;97(7):3491-6.

Diatchenko L, Lau YF, Campbell AP, Chenchik A, Moqadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Sverdlov ED, Siebert PD. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci U S A*. 1996 Jun 11;93(12):6025-30.

Dinel S, Bolduc C, Belleau P, Boivin A, Yoshioka M, Calvo E, Piedboeuf B, Snyder EE, Labrie F, St-Amand J. Reproducibility, bioinformatic analysis and power of the SAGE method to evaluate changes in transcriptome. *Nucleic Acids Res*. 2005;33(3):e26.

Dole VS, Myler PJ, Stuart KD, Madhubala R. Expression of biopterin transporter (BT1) protein in *Leishmania*. *FEMS Microbiol Lett*. 2002 Feb 19;208(1):89-91.

Dong H, Lin J, Han H, Jiang L, Zhao Q, Zhu S, Huang B. Analysis of differentially expressed genes in the precocious line of *Eimeria maxima* and its parent strain using suppression subtractive hybridization and cDNA microarrays. *Parasitol Res*. *In press*.

Dowse TJ, Pascall JC, Brown KD, Soldati D. Apicomplexan rhomboids have a potential role in microneme protein cleavage during host cell invasion. *Int J Parasitol*. 2005 Jun;35(7):747-56.

Dowse TJ, Soldati D. Rhomboid-like proteins in Apicomplexa: phylogeny and nomenclature. *Trends Parasitol*. 2005 Jun;21(6):254-8.

Dubremetz JF, Garcia-Reguet N, Conseil V, Fourmaux MN. Apical organelles and host-cell invasion by Apicomplexa. *Int J Parasitol*. 1998 Jul;28(7):1007-13.

Dufva M. *DNA Microarrays for Biomedical Research*. New York, USA: Humana Press; 2009. p. 1-22: Introduction to Microarray Technology.

Durham AM, Kashiwabara AY, Matsunaga FT, Ahagon PH, Rainone F, Varuzza L, Gruber A. EGene: a configurable pipeline generation system for automated sequence analysis. *Bioinformatics*. 2005 Jun 15;21(12):2812-3.

Echeverria PC, Matrajt M, Harb OS, Zappia MP, Costas MA, Roos DS, Dubremetz JF, Angel SO. *Toxoplasma gondii* Hsp90 is a potential drug target whose expression and subcellular localization are developmentally regulated. *J Mol Biol*. 2005 Jul 22;350(4):723-34.

Eckert J, Taylor M, Catchpole J, Licois D, Coudert P, Bucklar H. Identification of *Eimeria* species and strains: Morphological characteristics of oocysts. In: Eckert J, Braun R, Shirley MW and Coudert P, editors. Guidelines on techniques in coccidiosis research. Luxembourg: Office for official Publications of the European Communities; 1995.

Edgar SA. Sporulation of oocysts at specific temperatures and notes on the prepatent period of several species of avian coccidia. *J Parasitol.* 1955 Apr;41(2):214-6.

Edgar SA, Herrick CA, Fraser LA. Glycogen in the life cycle of coccidium, *Eimeria tenella*. *Trans. Amer. Microsc. Soc.* 1944;63:199-202.

Ellis J, Bumstead J. *Eimeria* species: studies using rRNA and rDNA probes. *Parasitology.* 1990 Aug;101 Pt 1:1-6.

Ellis J, Thurlby T. Changes in the messenger RNA population during sporulation of *Eimeria maxima*. *Parasitology.* 1991 Feb;102 Pt 1:1-8.

Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* 1998 Mar;8(3):175-85.

Ewing RM, Ben Kahla A, Poirot O, Lopez F, Audic S, Claverie JM. Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. *Genome Res.* 1999 Oct;9(10):950-9.

Fayer R. Epidemiology of Protozoan Infections - Coccidia. *Veterinary Parasitology.* 1980;6(1-3):75-103.

Fernandez-Gamba A, Leal MC, Morelli L, Castano EM. Insulin-degrading enzyme: structure-function relationship and its possible roles in health and disease. *Curr Pharm Des.* 2009;15(31):3644-55.

Fernandez S. Estudo das variações inter- e intra- específicas de *Eimeria* spp. de galinha doméstica através do polimorfismo de DNA amplificado randomicamente (RAPD). São Paulo: Universidade de São Paulo; 1999.

Fernandez S, Costa AC, Katsuyama AM, Madeira AM, Gruber A. A survey of the inter- and intraspecific RAPD markers of *Eimeria* spp. of the domestic fowl and the development of reliable diagnostic tools. *Parasitol Res.* 2003a Apr;89(6):437-45.

Fernandez S, Katsuyama AM, Kashiwabara AY, Madeira AM, Durham AM, Gruber A. Characterization of SCAR markers of *Eimeria* spp. of domestic fowl and construction of a public relational database (The *Eimeria* SCARdb). *FEMS Microbiol Lett.* 2004 Sep 1;238(1):183-8.

Fernandez S, Pagotto AH, Furtado MM, Katsuyama AM, Madeira AM, Gruber A. A multiplex PCR assay for the simultaneous detection and discrimination of the seven *Eimeria* species that infect domestic fowl. *Parasitology.* 2003b Oct;127(Pt 4):317-25.

Fernando MA. *Eimeria*: Infections of the Intestine. In: Long PL, editor. *Coccidiosis of Man and Domestic Animals*. Boston: CRC Press Inc; 1990. p. 63-75.

Ferro M. Desenvolvimento e validação de protocolos para a anotação automática de seqüências ORESTES de *Eimeria* spp. de galinha doméstica [Dissertação (Mestrado em Biologia da Relação Patógeno Hospedeiro)]. São Paulo: Instituto de Ciências Biomédicas da Universidade de São Paulo; 2008.

Fetterer RH, Barfield RC. Characterization of a developmentally regulated oocyst protein from *Eimeria tenella*. *J Parasitol.* 2003 Jun;89(3):553-64.

Fetterer RH, Jenkins MC, Miska KB, Barfield RC. Characterization of the antigen SO7 during development of *Eimeria tenella*. *J Parasitol.* 2007a Oct;93(5):1107-13.

Fetterer RH, Miska KB, Barfield RC. Partial purification and characterization of an aminopeptidase from *Eimeria tenella*. *J Parasitol.* 2005 Dec;91(6):1280-6.

Fetterer RH, Miska KB, Jenkins MC, Barfield RC. A conserved 19-kDa *Eimeria tenella* antigen is a profilin-like protein. *J Parasitol.* 2004 Dec;90(6):1321-8.

Fetterer RH, Miska KB, Jenkins MC, Barfield RC, Lillehoj H. Identification and characterization of a serpin from *Eimeria acervulina*. *J Parasitol.* 2008 Dec;94(6):1269-74.

Fetterer RH, Miska KB, Lillehoj H, Barfield RC. Serine protease activity in developmental stages of *Eimeria tenella*. *J Parasitol.* 2007b Apr;93(2):333-40.

Fietto JL, DeMarco R, Verjovski-Almeida S. Use of degenerate primers and touchdown PCR for construction of cDNA libraries. *Biotechniques.* 2002 Jun;32(6):1404-11.

Fitz-Coy SH. Antigenic variation among strains of *Eimeria maxima* and *E. tenella* of the chicken. *Avian Dis.* 1992 Jan-Mar;36(1):40-3.

Fleige T, Limenitakis J, Soldati-Favre D. Apicoplast: keep it or leave it. *Microbes Infect.* 2010 Apr;12(4):253-62.

Florent I, Charneau S, Grellier P. *Plasmodium falciparum* genes differentially expressed during merozoite morphogenesis. *Mol Biochem Parasitol.* 2004 May;135(1):143-8.

Fonseca RD, Carraro DM, Brentani H. Mining ORESTES no-match database: can we still contribute to cancer transcriptome? *Genetics and Molecular Research.* 2006;5(1):24-32.

Freeman M. Rhomboid proteases and their biological functions. *Annu Rev Genet.* 2008;42:191-210.

Fullwood MJ, Wei CL, Liu ET, Ruan Y. Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. *Genome Res.* 2009 Apr;19(4):521-32.

Garcia AM, Hernandez O, Aristizabal BH, De Souza Bernardes LA, Puccia R, Naranjo TW, Goldman GH, Goldman MH, Cano LE, Restrepo A, McEwen JG. Gene expression analysis of *Paracoccidioides brasiliensis* transition from conidium to yeast cell. *Med Mycol.* 2010 Feb;48(1):147-54.

Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Perteza M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature.* 2002 Oct 3;419(6906):498-511.

Gissot M, Briquet S, Refour P, Boschet C, Vaquero C. PfMyb1, a *Plasmodium falciparum* transcription factor, is required for intra-erythrocytic growth and controls key genes for cell cycle regulation. *J Mol Biol.* 2005 Feb 11;346(1):29-42.

Gomez C, Esther Ramirez M, Calixto-Galvez M, Medel O, Rodriguez MA. Regulation of gene expression in protozoa parasites. *J Biomed Biotechnol.* 2010;2010:726045.

Gopalakrishnan AM, Lopez-Estrano C. Comparative analysis of stage specific gene regulation of apicomplexan parasites: *Plasmodium falciparum* and *Toxoplasma gondii*. *Infect Disord Drug Targets*. 2010 Aug;10(4):303-11.

Graat EA, Henken AM, Ploeger HW, Noordhuizen JP, Vertommen MH. Rate and course of sporulation of oocysts of *Eimeria acervulina* under different environmental conditions. *Parasitology*. 1994 Jun;108(Pt 5):497-502.

Green CD, Simons JF, Taillon BE, Lewin DA. Open systems: panoramic views of gene expression. *J Immunol Methods*. 2001 Apr;250(1-2):67-79.

Gregory MW. Pathology of Coccidial Infections. In: Long PL, editor. *Coccidiosis of Man and Domestic Animals*. Boston: CRC Press; 1990. p. 235-62.

Grisard EC, Stoco PH, Wagner G, Sincero TC, Rotava G, Rodrigues JB, Snoeijer CQ, Koerich LB, Sperandio MM, Bayer-Santos E, Fragoso SP, Goldenberg S, Triana O, Vallejo GA, Tyler KM, Davila AM, Steindel M. Transcriptomic analyses of the avirulent protozoan parasite *Trypanosoma rangeli*. *Mol Biochem Parasitol*. 2010 Nov;174(1):18-25.

Gruber A. Expressed sequence tags. In: Dear PH, editor. *Bioinformatics*. Bloxham Mill, Oxfordshire, UK: Scion Publishing Limited; 2007. p. 141-67.

Hammond DM. Life Cycles and Development of Coccidia. In: Hammod DM and Long PL, editors. *The Coccidia*: Baltimore -University Park Press e London- Butterworths; 1973. p. 45-80.

Hammond DM, Long PL. *The Coccidia*. Baltimore -University Park Press London- Butterworths; 1973.

Han HY, Lin JJ, Zhao QP, Dong H, Jiang LL, Wang X, Han JF, Huang B. Construction of subtractive cDNA libraries of the sporogony stage of *Eimeria tenella* by suppression subtractive hybridization. *Sheng Wu Gong Cheng Xue Bao*. 2007;23(6):1005-10.

Han HY, Lin JJ, Zhao QP, Dong H, Jiang LL, Xu MQ, Zhu SH, Huang B. Identification of differentially expressed genes in early stages of *Eimeria tenella* by suppression subtractive hybridization and cDNA microarray. *J Parasitol*. 2010 2010;96(1):95-102.

Hanelt B, Lun CM, Adema CM. Comparative ORESTES-sampling of transcriptomes of immune-challenged *Biomphalaria glabrata* snails. *J Invertebr Pathol*. 2008 Oct;99(2):192-203.

Hoon M, Hayashizaki Y. Deep cap analysis gene expression (CAGE): genome-wide identification of promoters, quantification of their expression, and network inference. *Biotechniques*. 2008;44:627-32.

Hsieh JJ, Cheng EH, Korsmeyer SJ. Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell*. 2003 Oct 31;115(3):293-303.

Hu K, Johnson J, Florens L, Fraunholz M, Suravajjala S, DiLullo C, Yates J, Roos DS, Murray JM. Cytoskeletal components of an invasion machine--the apical complex of *Toxoplasma gondii*. *PLoS Pathog*. 2006 Feb;2(2):e13.

Huang X, Madan A. CAP3: A DNA sequence assembly program. *Genome Res*. 1999 Sep;9(9):868-77.

Huber R, Carrell RW. Implications of the 3-Dimensional Structure of Alpha-1-Antitrypsin for Structure and Function of Serpins. *Biochemistry*. 1989 Nov 14;28(23):8951-66.

Irving JA, Pike RN, Lesk AM, Whisstock JC. Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function. *Genome Res*. 2000 Dec;10(12):1845-64.

Jackson JB. Proton translocation by transhydrogenase. *FEBS Lett*. 2003 Nov 27;555(1):176-7.

Jansen G, Westerhof GR, Jarmuszewski MJ, Kathmann I, Rijksen G, Schornagel JH. Methotrexate transport in variant human CCRF-CEM leukemia cells with elevated levels of the reduced folate carrier. Selective effect on carrier-mediated transport of physiological concentrations of reduced folates. *J Biol Chem*. 1990 Oct 25;265(30):18272-7.

Jean L, Grosclaude J, Labbe M, Tomley F, Pery P. Differential localisation of an *Eimeria tenella* aspartyl proteinase during the infection process. *Int J Parasitol*. 2000 Sep;30(10):1099-107.

Jean L, Long M, Young J, Pery P, Tomley F. Aspartyl proteinase genes from apicomplexan parasites: evidence for evolution of the gene structure. *Trends Parasitol.* 2001a Oct;17(10):491-8.

Jean L, Pery P, Dunn P, Bumstead J, Billington K, Ryan R, Tomley F. Genomic organisation and developmentally regulated expression of an apicomplexan aspartyl proteinase. *Gene.* 2001b Jan 10;262(1-2):129-36.

Jones DT. THREADER: Protein Sequence Threading by Double Dynamic Programming. In: Salzberg S, Searls D and Kasif S, editors. *Computational Methods in Molecular Biology.* Amsterdam. Elsevier Science; 1998. p.285-312

Joyner LP. Host and Site Specificity. In: Long PL, editor. *The Biology of Coccidia.* Baltimore: University Park Press; 1982. p. 36-60.

Kall L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol.* 2004 May 14;338(5):1027-36.

Kang SW, Kweon CH, Lee EH, Choe SE, Jung SC, Quyen DV. The differentiation of transcription between tachyzoites and bradyzoites of in vitro cultured *Neospora caninum*. *Parasitol Res.* 2008a Oct;103(5):1011-8.

Kang SW, Lee EH, Jean YH, Choe SE, Van Quyen D, Lee MS. The differential protein expression profiles and immunogenicity of tachyzoites and bradyzoites of in vitro cultured *Neospora caninum*. *Parasitol Res.* 2008b Sep;103(4):905-13.

Kawahara F, Taira K, Nagai S, Onaga H, Onuma M, Nunoya T. Detection of five avian *Eimeria* species by species-specific real-time polymerase chain reaction assay. *Avian Dis.* 2008 Dec;52(4):652-6.

Kim YC, Jung YC, Xuan Z, Dong H, Zhang MQ, Wang SM. Pan-genome isolation of low abundance transcripts using SAGE tag. *FEBS Lett.* 2006 Dec 11;580(28-29):6721-9.

Kinnaird JH, Bumstead JM, Mann DJ, Ryan R, Shirley MW, Shiels BR, Tomley FM. EtCRK2, a cyclin-dependent kinase gene expressed during the sexual and asexual phases of the *Eimeria tenella* life cycle. *Int J Parasitol.* 2004 May;34(6):683-92.

Klotz C, Marhofer RJ, Selzer PM, Lucius R, Pogonka T. *Eimeria tenella*: identification of secretory and surface proteins from expressed sequence tags. *Exp Parasitol*. 2005 Sep;111(1):14-23.

Kramer RA, Tomchak LA, McAndrew SJ, Becker K, Hug D, Pasamontes L, Humbelin M. An *Eimeria tenella* gene encoding a protein with homology to the nucleotide transhydrogenases of *Escherichia coli* and bovine mitochondria. *Mol Biochem Parasitol*. 1993 Aug;60(2):327-31.

Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*. 2001 Jan 19;305(3):567-80.

Kronegg J, Buloz D. Detection/prediction of GPI cleavage site (GPI-anchor) in a protein (DGPI). 1999. Available from: <http://129.194.185.165/dgpi/> [2007 Jan 25].

Laurent F, Bourdieu C, Kaga M, Chilmonczyk S, Zgrzebski G, Yvore P, Pery P. Cloning and characterization of an *Eimeria acervulina* sporozoite gene homologous to aspartyl proteinases. *Mol Biochem Parasitol*. 1993 Dec;62(2):303-12.

Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, Haynes JD, De La Vega P, Holder AA, Batalov S, Carucci DJ, Winzeler EA. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science*. 2003 Sep 12;301(5639):1503-8.

Leerkes MR, Caballero OL, Mackay A, Torloni H, O'Hare MJ, Simpson AJ, de Souza SJ. In silico comparison of the transcriptome derived from purified normal breast cells and breast tumor cell lines reveals candidate upregulated genes in breast tumor cells. *Genomics*. 2002 Feb;79(2):257-65.

Levine ND. *Protozoan Parasites of Domestic Animals and of Man*. Minneapolis: Bugess Publishing Company; 1961. p. 432: The Telosporasida and the Coccidia Proper.

Levine ND. Introduction, History, and Taxonomy. In: Hammod DM and Long PL, editors. *The Coccidia*: Baltimore -University Park Press; 1973. p. 1-22.

Levine ND. Progress in taxonomy of the Apicomplexan protozoa. *J Protozool*. 1988 Nov;35(4):518-20.

Levine PP. A new coccidium pathogenic for chickens. *Eimeria brunetti* n. sp. (Protozoa: Eimeriidae). *Cornell Veterinarian* 1942;32:430-9.

Lew AE, Anderson GR, Minchin CM, Jeston PJ, Jorgensen WK. Inter- and intra-strain variation and PCR detection of the internal transcribed spacer 1 (ITS-1) sequences of Australian isolates of *Eimeria* species from chickens. *Veterinary Parasitology*. 2003 Feb 28;112(1-2):33-50.

Li H, Han Z, Lu Y, Lin Y, Zhang L, Wu Y, Wang H. Isolation and functional characterization of a dynamin-like gene from *Plasmodium falciparum*. *Biochem Biophys Res Commun*. 2004a Jul 30;320(3):664-71.

Li L, Brunk BP, Kissinger JC, Pape D, Tang K, Cole RH, Martin J, Wylie T, Dante M, Fogarty SJ, Howe DK, Liberator P, Diaz C, Anderson J, White M, Jerome ME, Johnson EA, Radke JA, Stoeckert CJ, Jr., Waterston RH, Clifton SW, Roos DS, Sibley LD. Gene discovery in the apicomplexa as revealed by EST sequencing and assembly of a comparative gene database. *Genome Res*. 2003 Mar;13(3):443-54.

Li L, Crabtree J, Fischer S, Pinney D, Stoeckert CJ, Jr., Sibley LD, Roos DS. ApiEST-DB: analyzing clustered EST data of the apicomplexan parasites. *Nucleic Acids Res*. 2004b Jan 1;32:D326-8.

Liang P, Pardee AB. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*. 1992 Aug 14;257(5072):967-71.

Liberator P, Anderson J, Feiglin M, Sardana M, Griffin P, Schmatz D, Myers RW. Molecular cloning and functional expression of mannitol-1-phosphatase from the apicomplexan parasite *Eimeria tenella*. *J Biol Chem*. 1998 Feb 13;273(7):4237-44.

Ling KH, Rajandream MA, Rivailler P, Ivens A, Yap SJ, Madeira AM, Mungall K, Billington K, Yee WY, Bankier AT, Carroll F, Durham AM, Peters N, Loo SS, Isa MN, Novaes J, Quail M, Rosli R, Nor Shamsudin M, Sobreira TJ, Tivey AR, Wai SF, White S, Wu X, Kerhornou A, Blake D, Mohamed R, Shirley M, Gruber A, Berriman M, Tomley F, Dear PH, Wan KL. Sequencing and analysis of chromosome 1 of *Eimeria tenella* reveals a unique segmental organization. *Genome Res*. 2007 Mar;17(3):311-9.

Liu ET. Genomic technologies and the interrogation of the transcriptome. *Mech Ageing Dev*. 2005 Jan;126(1):153-9.

Llinas M, DeRisi JL. Pernicious plans revealed: *Plasmodium falciparum* genome wide expression analysis. *Curr Opin Microbiol*. 2004 Aug;7(4):382-7.

Lockyer AE, Spinks J, Kane RA, Hoffmann KF, Fitzpatrick JM, Rollinson D, Noble LR, Jones CS. *Biomphalaria glabrata* transcriptome: cDNA microarray profiling identifies resistant- and susceptible-specific gene expression in haemocytes from snail strains exposed to *Schistosoma mansoni*. *BMC Genomics*. 2008;9:634.

Lockyer AE, Spinks JN, Walker AJ, Kane RA, Noble LR, Rollinson D, Dias-Neto E, Jones CS. *Biomphalaria glabrata* transcriptome: identification of cell-signalling, transcriptional control and immune-related genes from open reading frame expressed sequence tags (ORESTES). *Dev Comp Immunol*. 2007;31(8):763-82.

Long PL. A study of *Eimeria maxima* Tyzzer, 1929, a coccidium of the fowl (*Gallus gallus*). *Ann Trop Med Parasitol*. 1959 Sep;53:325-33.

Long PL, Joyner LP. Problems in the identification of species of *Eimeria*. *J Protozool*. 1984 Nov;31(4):535-41.

Long PL, Millard BJ. Immunological differences in *Eimeria maxima*: effect of a mixed immunizing inoculum on heterologous challenge. *Parasitology*. 1979 Dec;79(3):451-7.

Long PL, Millard BJ, Joyner LP, Norton CC. A guide to laboratory techniques used in the study and diagnosis of avian coccidiosis. *Folia Vet Lat*. 1976 Jul-Sep;6(3):201-17.

Lopez-Estrano C. Comparative analysis of stage specific gene regulation of apicomplexan parasites: *Plasmodium falciparum* and *Toxoplasma gondii*. *Infect Disord Drug Targets*. 2010 Aug;10(4):240-1.

Mai K, Sharman PA, Walker RA, Katrib M, De Souza D, McConville MJ, Wallach MG, Belli SI, Ferguson DJ, Smith NC. Oocyst wall formation and composition in coccidian parasites. *Mem Inst Oswaldo Cruz*. 2009 Mar;104(2):281-9.

Maia RM, Valente V, Cunha MA, Sousa JF, Araujo DD, Silva WA, Jr., Zago MA, Dias-Neto E, Souza SJ, Simpson AJ, Monesi N, Ramos RG, Espreafico EM, Paco-Larson ML. Identification of unannotated exons of low abundance transcripts in *Drosophila melanogaster* and cloning of a new serine protease gene upregulated upon injury. *BMC Genomics*. 2007;8:249.

Marchler-Bauer A, Anderson JB, Derbyshire MK, DeWeese-Scott C, Gonzales NR, Gwadz M, Hao L, He S, Hurwitz DI, Jackson JD, Ke Z, Krylov D, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Thanki N, Yamashita RA, Yin JJ, Zhang D, Bryant SH. CDD: a conserved domain database for interactive domain family analysis. *Nucleic Acids Res.* 2007 Jan;35:D237-40.

Marchler-Bauer A, Panchenko AR, Shoemaker BA, Thiessen PA, Geer LY, Bryant SH. CDD: a database of conserved domain alignments with links to domain three-dimensional structure. *Nucleic Acids Res.* 2002 Jan 1;30(1):281-3.

Maresca B, Carratu L. The biology of the heat shock response in parasites. *Parasitol Today.* 1992 Aug;8(8):260-6.

Martin AG, Danforth HD, Barta JR, Fernando MA. Analysis of immunological cross-protection and sensitivities to anticoccidial drugs among five geographical and temporal strains of *Eimeria maxima*. *Int J Parasitol.* 1997 May;27(5):527-33.

Massimine KM, Doan LT, Atreya CA, Stedman TT, Anderson KS, Joiner KA, Coppens I. *Toxoplasma gondii* is capable of exogenous folate transport. A likely expansion of the BT1 family of transmembrane proteins. *Mol Biochem Parasitol.* 2005 Nov;144(1):44-54.

Mathis GF, McDougald LR, McMurray B. Effectiveness of therapeutic anticoccidial drugs against recently isolated coccidia. *Poult Sci.* 1984 Jun;63(6):1149-53.

McDonald V, Shirley MW. Past and future: vaccination against *Eimeria*. *Parasitology.* 2009 Oct;136(12):1477-89.

McDougald LR, Reid WM. Coccidiosis. In: Calnek BW, editor. *Diseases of Poultry*. 9th ed. Ames: Iowa State University Press; 1995. p. 929.

Mello BP, Abrantes EF, Torres CH, Machado-Lima A, Fonseca Rda S, Carraro DM, Brentani RR, Reis LF, Brentani H. No-match ORESTES explored as tumor markers. *Nucleic Acids Res.* 2009 May;37(8):2607-17.

Min W, Dalloul RA, Lillehoj HS. Application of biotechnological tools for coccidia vaccine development. *J Vet Sci.* 2004 Dec;5(4):279-88.

Miska KB, Fetterer RH, Barfield RC. Analysis of transcripts expressed by *Eimeria tenella* oocysts using subtractive hybridization methods. J Parasitol. 2004 Dec;90(6):1245-52.

Miska KB, Fetterer RH, Min W, Lillehoj HS. Heat shock protein 90 genes of two species of poultry *Eimeria*: expression and evolutionary analysis. J Parasitol. 2005 Apr;91(2):300-6.

Miska KB, Fetterer RH, Rosenberg GH. Analysis of transcripts from intracellular stages of *Eimeria acervulina* using expressed sequence tags. J Parasitol. 2008 Apr;94(2):462-6.

Monné L, Hönig G. On the properties of the shells of the coccidian oocysts. Arkiv för Zoologi. 1954 12 may 1954;7(15):251-6.

Moody DE. Genomics Techniques: An overview of methods for the study of gene expression. J Anim Sci. 2001;79:E128-E35.

Morozova O, Hirst M, Marra MA. Applications of New Sequencing Technologies for Transcriptome Analysis. Annual Review of Genomics and Human Genetics. 2009;10:135-51.

Morrisette NS, Sibley LD. Cytoskeleton of apicomplexan parasites. Microbiol Mol Biol Rev. 2002 Mar;66(1):21-38.

Morse DL, Carroll D, Weberg L, Borgstrom MC, Ranger-Moore J, Gillies RJ. Determining suitable internal standards for mRNA quantification of increasing cancer progression in human breast cells by real-time reverse transcriptase polymerase chain reaction. Anal Biochem. 2005 Jul 1;342(1):69-77.

Mulder N, Apweiler R. InterPro and InterProScan: tools for protein sequence classification and comparison. Methods Mol Biol. 2007;396:59-70.

Mullapudi N, Joseph SJ, Kissinger JC. Identification and functional characterization of cis-regulatory elements in the apicomplexan parasite *Toxoplasma gondii*. Genome Biol. 2009;10(4):R34.

Munasinghe A, Patankar S, Cook BP, Madden SL, Martin RK, Kyle DE, Shoaibi A, Cummings LM, Wirth DF. Serial analysis of gene expression (SAGE) in *Plasmodium*

falciparum: application of the technique to A-T rich genomes. Mol Biochem Parasitol. 2001 Mar;113(1):23-34.

Nagai MA, Da Ros N, Neto MM, Junior SRD, Brentani MM, Hirata R, Neves EJ. Gene expression profiles in breast tumors regarding the presence or absence of estrogen and progesterone receptors. International Journal of Cancer. 2004 Oct 10;111(6):892-9.

Nagaraj SH, Gasser RB, Ranganathan S. A hitchhiker's guide to expressed sequence tag (EST) analysis. Brief Bioinform. 2007 Jan;8(1):6-21.

Nathan DF, Vos MH, Lindquist S. In vivo functions of the *Saccharomyces cerevisiae* Hsp90 chaperone. Proc Natl Acad Sci U S A. 1997 Nov 25;94(24):12949-56.

Newport G, Culpepper J, Agabian N. Parasite Heat-Shock Proteins. Parasitology Today. 1988 Nov;4(11):306-12.

Ng P, Wei CL, Ruan Y. Paired-end diTagging for transcriptome and genome analysis. Curr Protoc Mol Biol. 2007 Jul;Chapter 21:Unit 21 12.

Ng ST, Sanusi Jangi M, Shirley MW, Tomley FM, Wan KL. Comparative EST analyses provide insights into gene expression in two asexual developmental stages of *Eimeria tenella*. Exp Parasitol. 2002 Jun-Jul;101(2-3):168-73.

Norton CC, Chard MJ. The oocyst sporulation time of *Eimeria* species from the fowl. Parasitology. 1983 Apr;86(Pt 2):193-8.

Novaes J. Análise da expressão diferencial entre merozoítos e esporozoítos de *Eimeria tenella* empregando a técnica de LongSAGE. São Paulo: Universidade de São Paulo; 2009.

Nunes FM, Valente V, Sousa JF, Cunha MA, Pinheiro DG, Maia RM, Araujo DD, Costa MC, Martins WK, Carvalho AF, Monesi N, Nascimento AM, Peixoto PM, Silva MF, Ramos RG, Reis LF, Dias-Neto E, Souza SJ, Simpson AJ, Zago MA, Soares AE, Bitondi MM, Espreafico EM, Espindola FS, Paco-Larson ML, Simoes ZL, Hartfelder K, Silva WA, Jr. The use of Open Reading frame ESTs (ORESTES) for analysis of the honey bee transcriptome. BMC Genomics. 2004 Nov 3;5(1):84.

Okhuysen PC, Chappell CL, Kettner C, Sterling CR. Cryptosporidium parvum metalloaminopeptidase inhibitors prevent in vitro excystation. Antimicrob Agents Chemother. 1996 Dec;40(12):2781-4.

Padda RS, Tsai A, Chappell CL, Okhuysen PC. Molecular cloning and analysis of the Cryptosporidium parvum aminopeptidase N gene. Int J Parasitol. 2002 Feb;32(2):187-97.

Pease AC, Solas D, Sullivan EJ, Cronin MT, Holmes CP, Fodor SP. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. Proc Natl Acad Sci U S A. 1994 May 24;91(11):5022-6.

Peirson SN, Butler JN, Foster RG. Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. Nucleic Acids Res. 2003 Jul 15;31(14):e73.

Péroval M, Péry P, Labbé M. The heat shock protein 90 of *Eimeria tenella* is essential for invasion of host cell and schizont growth. Int J Parasitol. 2006 Sep;36(10-11):1205-15.

Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 2001 May 1;29(9):e45.

Pfaffl MW. Relative quantification. In: Dorak MT, editor. Real-time PCR. New York, USA: Taylor & Francis Group; 2006. p. 63-82.

Pittilo RM, Ball SJ. The ultrastructural development of the oocyst wall of *Eimeria maxima*. Parasitology. 1980 Aug;81(1):115-22.

Pittilo RM, Ball SJ. Ultrastructural observations on the sporogony of *Eimeria maxima*. Int J Parasitol. 1985 Dec;15(6):617-20.

Ponchel F. Real-time PCR using SYBR® Green. In: Dorak MT, editor. Real-time PCR. New York, USA: Taylor & Francis Group; 2006. p. 139-54.

Possenti A, Cherchi S, Bertuccini L, Pozio E, Dubey JP, Spano F. Molecular characterisation of a novel family of cysteine-rich proteins of *Toxoplasma gondii* and ultrastructural evidence of oocyst wall localisation. Int J Parasitol. 2010 Dec;40(14):1639-49.

Praefcke GJ, McMahon HT. The dynamin superfamily: universal membrane tubulation and fission molecules? *Nat Rev Mol Cell Biol.* 2004 Feb;5(2):133-47.

Procnier JD, Fernando MA, Barta JR. Species and strain differentiation of *Eimeria* spp. of the domestic fowl using DNA polymorphisms amplified by arbitrary primers. *Parasitol Res.* 1993;79(2):98-102.

Pszenny V, Angel SO, Duschak VG, Paulino M, Ledesma B, Yabo MI, Guarnera E, Ruiz AM, Bontempi EJ. Molecular cloning, sequencing and expression of a serine proteinase inhibitor gene from *Toxoplasma gondii*. *Mol Biochem Parasitol.* 2000 Apr 15;107(2):241-9.

Pszenny V, Ledesma BE, Matrajt M, Duschak VG, Bontempi EJ, Dubremetz JF, Angel SO. Subcellular localization and post-secretory targeting of TgPI, a serine proteinase inhibitor from *Toxoplasma gondii*. *Mol Biochem Parasitol.* 2002 May;121(2):283-6.

Radke JR, Behnke MS, Mackey AJ, Radke JB, Roos DS, White MW. The transcriptome of *Toxoplasma gondii*. *BMC Biol.* 2005;3:26.

Radvanyi L, Singh-Sandhu D, Gallichan S, Lovitt C, Pedyczak A, Mallo G, Gish K, Kwok K, Hanna W, Zubovits J, Armes J, Venter D, Hakimi J, Shortreed J, Donovan M, Parrington M, Dunn P, Oomen R, Tartaglia J, Berinstein NL. The gene associated with trichorhinophalangeal syndrome in humans is overexpressed in breast cancer. *Proc Natl Acad Sci U S A.* 2005 Aug 2;102(31):11005-10.

Refega S, Girard-Misguich F, Bourdieu C, Pery P, Labbe M. Gene discovery in *Eimeria tenella* by immunoscreening cDNA expression libraries of sporozoites and schizonts with chicken intestinal antibodies. *Vet Parasitol.* 2003 Apr 2;113(1):19-33.

Rivals E, Boureux A, Lejeune M, Ottones F, Perez OP, Tarhio J, Pierrat F, Ruffle F, Commes T, Marti J. Transcriptome annotation using tandem SAGE tags. *Nucleic Acids Res.* 2007;35(17):e108.

Romano CM. Caracterização molecular e análise comparativa de genomas mitocondriais de *Eimeria* spp. de galinha doméstica [Dissertação (Mestrado em Biologia da Relação Patógeno-Hospedeiro)]. São Paulo: Instituto de Ciências Biomédicas da Universidade de São Paulo; 2004.

Rose ME, Long PL. Immunity to four species of *Eimeria* in fowls. *Immunology.* 1962 Jan;5:79-92.

Ryan R, Shirley M, Tomley F. Mapping and expression of microneme genes in *Eimeria tenella*. *Int J Parasitol*. 2000 Dec;30(14):1493-9.

Ryley JF. Cytochemistry, Physiology and Biochemistry. In: Hammond DM and Long PL, editors. *The Coccidia: Eimeria, Isospora, Toxoplasma and Related Genera*. London: University Park Press; 1973. p. 145-82.

Ryley JF, Bentley M, Manners DJ, Stark JR. Amylopectin, the storage polysaccharide of the Coccidia *Eimeria brunetti* and *E. tenella*. *J Parasitol*. 1969 Aug;55(4):839-45.

Said HM, Chatterjee N, Haq RU, Subramanian VS, Ortiz A, Matherly LH, Sirotiak FM, Halsted C, Rubin SA. Adaptive regulation of intestinal folate uptake: effect of dietary folate deficiency. *Am J Physiol Cell Physiol*. 2000 Dec;279(6):C1889-95.

Sakabe NJ, de Souza JE, Galante PA, de Oliveira PS, Passetti F, Brentani H, Osorio EC, Zaiats AC, Leerkes MR, Kitajima JP, Brentani RR, Strausberg RL, Simpson AJ, de Souza SJ. ORESTES are enriched in rare exon usage variants affecting the encoded proteins. *C R Biol*. 2003 Oct-Nov;326(10-11):979-85.

Sambrook J, Russell DW. *Molecular Cloning - A Laboratory Manual*. 3rd ed. New York: Cold Spring Harbor Laboratory Press; 2001.

Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. 1977 Dec;74(12):5463-7.

Santos JM, Lebrun M, Daher W, Soldati D, Dubremetz JF. Apicomplexan cytoskeleton and motors: key regulators in morphogenesis, cell division, transport and motility. *Int J Parasitol*. 2009 Jan;39(2):153-62.

Sargent TD, Dawid IB. Differential Gene-Expression in the Gastrula of *Xenopus-Laewis*. *Science*. 1983;222(4620):133-9.

Sarkis AS. Global analysis of differential gene expression between prostate cancer and normal prostate tissues using cDNA microarray of open-reading frame expressed sequence tags (ORESTES). *Prostate Cancer Prostatic Dis*. 2000 Dec;3(S1):S36.

Schaap D, Arts G, Kroeze J, Niessen R, Roosmalen-Vos SV, Spreeuwenberg K, Kuiper CM, Beek-Verhoeven NV, Kok JJ, Knegt RM, Vermeulen AN. An *Eimeria*

vaccine candidate appears to be lactate dehydrogenase; characterization and comparative analysis. *Parasitology*. 2004 Jun;128(Pt 6):603-16.

Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*. 1995 Oct 20;270(5235):467-70.

Schmatz DM. The mannitol cycle in *Eimeria*. *Parasitology*. 1997;114(Suppl):S81-9.

Schnitzler BE, Thebo PL, Mattsson JG, Tomley FM, Shirley MW. Development of a diagnostic PCR assay for the detection and discrimination of four pathogenic *Eimeria* species of the chicken. *Avian Pathol*. 1998;27(5):490-7.

Schnitzler BE, Thebo PL, Tomley FM, Uggla A, Shirley MW. PCR identification of chicken *Eimeria*: a simplified read-out. *Avian Pathol*. 1999 Feb;28(1):89-93.

Schwarz RS, Fetterer RH, Rosenberg GH, Miska KB. Coccidian merozoite transcriptome analysis from *Eimeria maxima* in comparison to *Eimeria tenella* and *Eimeria acervulina*. *J Parasitol*. Feb;96(1):49-57.

Schwarz RS, Fetterer RH, Rosenberg GH, Miska KB. Coccidian Merozoite Transcriptome Analysis from *Eimeria Maxima* in Comparison to *Eimeria Tenella* and *Eimeria Acervulina*. *Journal of Parasitology*. 2010 Feb;96(1):49-57.

Schwarz RS, Jenkins MC, Klopp S, Miska KB. Genomic Analysis of *Eimeria* Spp. Populations in Relation to Performance Levels of Broiler Chicken Farms in Arkansas and North Carolina. *Journal of Parasitology*. 2009 Aug;95(4):871-80.

Shah T, de Villiers E, Nene V, Hass B, Taracha E, Gardner MJ, Sansom C, Pelle R, Bishop R. Using the transcriptome to annotate the genome revisited: application of massively parallel signature sequencing (MPSS). *Gene*. 2006 Jan 17;366(1):104-8.

Shiraki T, Kondo S, Katayama S, Waki K, Kasukawa T, Kawaji H, Kodzius R, Watahiki A, Nakamura M, Arakawa T, Fukuda S, Sasaki D, Podhajaska A, Harbers M, Kawai J, Carninci P, Hayashizaki Y. Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. *Proc Natl Acad Sci U S A*. 2003 Dec 23;100(26):15776-81.

Shirley MW. Enzyme variation in *Eimeria* species of the chicken. *Parasitology*. 1975 Dec;71(3):369-76.

Shirley MW. Electrophoretic variation of enzymes: a further marker for genetic studies of the *Eimeria*. *Z Parasitenkd.* 1978 Sep 4;57(1):83-7.

Shirley MW. The genome of *Eimeria tenella*: further studies on its molecular organisation. *Parasitol Res.* 1994;80(5):366-73.

Shirley MW. *Eimeria* species and strains of chicken. In: Eckert J, Braun R, Shirley MW and Coudert P, editors. COST 89/820 Biotechnology: Guidelines on techniques in coccidiosis research. Brussels: Luxembourg European Commission; 1995. p. 1-24.

Shirley MW. *Eimeria* spp. from the chicken: occurrence, identification and genetics. *Acta Vet Hung.* 1997;45(3):331-47.

Shirley MW. The genome of *Eimeria* spp., with special reference to *Eimeria tenella*--a coccidium from the chicken. *Int J Parasitol.* 2000 Apr 10;30(4):485-93.

Shirley MW, Blake D, White SE, Sheriff R, Smith AL. Integrating genetics and genomics to identify new leads for the control of *Eimeria* spp. *Parasitology.* 2004a;128(Suppl 1):S33-42.

Shirley MW, Harvey DA. *Eimeria tenella*: infection with a single sporocyst gives a clonal population. *Parasitology.* 1996 Jun;112(Pt 6):523-8.

Shirley MW, Harvey DA. A genetic linkage map of the apicomplexan protozoan parasite *Eimeria tenella*. *Genome Res.* 2000 Oct;10(10):1587-93.

Shirley MW, Ivens A, Gruber A, Madeira AM, Wan KL, Dear PH, Tomley FM. The *Eimeria* genome projects: a sequence of events. *Trends Parasitol.* 2004b May;20(5):199-201.

Shirley MW, Smith AL, Blake DP. Challenges in the successful control of the avian coccidia. *Vaccine.* 2007 Jul 26;25(30):5540-7.

Shirley MW, Smith AL, Tomley FM. The biology of avian *Eimeria* with an emphasis on their control by vaccination. *Adv Parasitol.* 2005;60:285-330.

Silveira NJ, Varuzza L, Machado-Lima A, Lauretto MS, Pinheiro DG, Rodrigues RV, Severino P, Nobrega FG, Silva WA, Jr., de BPCA, Tajara EH. Searching for

molecular markers in head and neck squamous cell carcinomas (HNSCC) by statistical and bioinformatic analysis of larynx-derived SAGE libraries. *BMC Med Genomics*. 2008;1:56.

Smith CK, Lee DE. Monosaccharide transport by *Eimeria tenella* sporozoites. *J Parasitol*. 1986 Feb;72(1):163-9.

Smith NC, Wallach M, Miller CM, Braun R, Eckert J. Maternal transmission of immunity to *Eimeria maxima*: western blot analysis of protective antibodies induced by infection. *Infect Immun*. 1994 Nov;62(11):4811-7.

Snoeijer CQ, Picchi GF, Dambros BP, Steindel M, Goldenberg S, Fragoso SP, Lorenzini DM, Grisard EC. Trypanosoma rangeli Transcriptome Project: Generation and analysis of expressed sequence tags. *Kinetoplastid Biol Dis*. 2004 May 13;3(1):1.

Sterky F, Lundeberg J. Sequence analysis of genes and genomes. *J Biotechnol*. 2000 Jan 7;76(1):1-31.

Stotish RL, Wang CC, Meyenhofer M. Structure and composition of the oocyst wall of *Eimeria tenella*. *J Parasitol*. 1978 Dec;64(6):1074-81.

Stucki U, Braun R, Roditi I. *Eimeria tenella*: characterization of a 5S ribosomal RNA repeat unit and its use as a species-specific probe. *Exp Parasitol*. 1993 Feb;76(1):68-75.

Sutton CA, Shirley MW, Wisher MH. Characterization of coccidial proteins by two-dimensional sodium dodecyl sulphate-polyacrylamide gel electrophoresis. *Parasitology*. 1989 Oct;99(Pt 2):175-87.

Tang J, Wong RN. Evolution in the structure and function of aspartic proteases. *J Cell Biochem*. 1987 Jan;33(1):53-63.

Tomley FM, Billington KJ, Bumstead JM, Clark JD, Monaghan P. EtMIC4: a microneme protein from *Eimeria tenella* that contains tandem arrays of epidermal growth factor-like repeats and thrombospondin type-I repeats. *Int J Parasitol*. 2001 Oct;31(12):1303-10.

Tomley FM, Bumstead JM, Billington KJ, Dunn PP. Molecular cloning and characterization of a novel acidic microneme protein (Etmic-2) from the apicomplexan protozoan parasite, *Eimeria tenella*. *Mol Biochem Parasitol*. 1996 Aug;79(2):195-206.

Tomley FM, Soldati DS. Mix and match modules: structure and function of microneme proteins in apicomplexan parasites. *Trends Parasitol*. 2001 Feb;17(2):81-8.

Toulza E, Mattiuzzo NR, Galliano MF, Jonca N, Dossat C, Jacob D, de Daruvar A, Wincker P, Serre G, Guerrin M. Large-scale identification of human genes implicated in epidermal barrier function. *Genome Biol*. 2007;8(6):R107.

Tuda J, Mongan AE, Tolba ME, Imada M, Yamagishi J, Xuan X, Wakaguri H, Sugano S, Sugimoto C, Suzuki Y. Full-parasites: database of full-length cDNAs of apicomplexa parasites, 2010 update. *Nucleic Acids Res*. 2010 Nov 4:1-7.

Tuteja R, Tuteja N. Serial analysis of gene expression (SAGE): application in cancer research. *Med Sci Monit*. 2004 Jun;10(6):RA132-40.

Tyzzar EE. Coccidiosis in gallinaceous birds. *Amer J Hyg*. 1929(10):269-83.

Urban S, Lee JR, Freeman M. *Drosophila* rhomboid-1 defines a family of putative intramembrane serine proteases. *Cell*. 2001 Oct 19;107(2):173-82.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 2002 Jun 18;3(7):RESEARCH0034.

Varuzza L. Métodos Estatísticos para a análise de bibliotecas digitais de expressão gênica [Doutorado (Programa Interunidades de Bioinformática)]. São Paulo: Universidade de São Paulo; 2008.

Varuzza L, Pereira CAB. Significance test for comparing digital gene expression profiles: Partial likelihood application. *Chilean Journal of Statistics*. 2010;1(1):91-102.

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. *Science*. 1995 Oct 20;270(5235):484-7.

Vencio RZ, Varuzza L, de BPCA, Brentani H, Shmulevich I. Simcluster: clustering enumeration gene expression data on the simplex space. *BMC Bioinformatics*. 2007;8:246.

Verjovski-Almeida S, DeMarco R, Martins EA, Guimaraes PE, Ojopi EP, Paquola AC, Piazza JP, Nishiyama MY, Jr., Kitajima JP, Adamson RE, Ashton PD, Bonaldo MF, Coulson PS, Dillon GP, Farias LP, Gregorio SP, Ho PL, Leite RA, Malaquias LC, Marques RC, Miyasato PA, Nascimento AL, Ohlweiler FP, Reis EM, Ribeiro MA, Sa RG, Stukart GC, Soares MB, Gargioni C, Kawano T, Rodrigues V, Madeira AM, Wilson RA, Menck CF, Setubal JC, Leite LC, Dias-Neto E. Transcriptome analysis of the acelomate human parasite *Schistosoma mansoni*. *Nat Genet*. 2003 Oct;35(2):148-57.

Vermeulen AN, Kok JJ, van den Boogaart P, Dijkema R, Claessens JA. Eimeria refractile body proteins contain two potentially functional characteristics: transhydrogenase and carbohydrate transport. *FEMS Microbiol Lett*. 1993 Jun 15;110(2):223-9.

Vetterling JM, Doran DJ. Storage polysaccharide in coccidial sporozites after excystation and penetration of cells. *J Protozool*. 1969 Nov;16(4):772-5.

Vrba V, Blake DP, Poplstein M. Quantitative real-time PCR assays for detection and quantification of all seven Eimeria species that infect the chicken. *Vet Parasitol*. 2010 Sep 15; 174(3-4):183-90

Wagenbach GE, Burns WC. Structure and respiration of sporulating Eimeria stiedae and E. tenella oocysts. *J Protozool*. 1969 May;16(2):257-63.

Waldenstedt L, Elwinger K, Lunden A, Thebo P, Uggla A. Sporulation of Eimeria maxima oocysts in litter with different moisture contents. *Poult Sci*. 2001 Oct;80(10):1412-5.

Wan KL, Chong SP, Ng ST, Shirley MW, Tomley FM, Jangi MS. A survey of genes in *Eimeria tenella* merozoites by EST sequencing. *Int J Parasitol*. 1999 Dec;29(12):1885-92.

Wang CC. Biochemistry and Physiology of Coccidia. In: Long P, editor. *The biology of the Coccidia*. Georgia: Edward Arnold; 1982. p. 168-228.

Wang CC, Stotish RL. Changes of Nucleic-Acids and Proteins in Oocysts of Eimeria-Tenella during Sporulation. *Journal of Protozoology*. 1975;22(3):438-43.

Wang CC, Weppelman RM, Lopez-Ramos B. Isolation of amylopectin granules and identification of amylopectin phosphorylase in the oocysts of *Eimeria tenella*. *J Protozool*. 1975 Nov;22(4):560-4.

Wang J, Zhang W, Chen T, Wang M. *Eimeria tenella* cDNA library construction and expressed sequence tags analysis. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi*. 2007;24(6):1357-62.

Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009 Jan;10(1):57-63.

Wastling JM, Xia D, Sohal A, Chaussepied M, Pain A, Langsley G. Proteomes and transcriptomes of the Apicomplexa--where's the message? *Int J Parasitol*. 2009 Jan;39(2):135-43.

Webster WA, McFadden GI. Organelle division: dynamin-related proteins in apicomplexans. *Curr Biol*. 2009 Apr 28;19(8):R334-6.

Weiss LM, Fiser A, Angeletti RH, Kim K. *Toxoplasma gondii* proteomics. *Expert Rev Proteomics*. 2009 Jun;6(3):303-13.

Welsh J, Chada K, Dalal SS, Cheng R, Ralph D, McClelland M. Arbitrarily primed PCR fingerprinting of RNA. *Nucleic Acids Res*. 1992 Oct 11;20(19):4965-70.

Williams RB. Epidemiological aspects of the use of live anticoccidial vaccines for chickens. *Int J Parasitol*. 1998 Jul;28(7):1089-98.

Wilson PA, Fairbairn D. Biochemistry of Sporulation in Oocysts of *Eimeria Acervulina*. *Journal of Protozoology*. 1961;8(4):410-6.

Wilson RJ, Denny PW, Preiser PR, Rangachari K, Roberts K, Roy A, Whyte A, Strath M, Moore DJ, Moore PW, Williamson DH. Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. *J Mol Biol*. 1996 Aug 16;261(2):155-72.

Witcombe DM, Belli SI, Wallach MG, Smith NC. Molecular characterisation of EmTFP250: a novel member of the TRAP protein family in *Eimeria maxima*. *Int J Parasitol*. 2003 Jul;33(7):691-702.

Xia D, Sanderson SJ, Jones AR, Prieto JH, Yates JR, Bromley E, Tomley FM, Lal K, Sinden RE, Brunk BP, Roos DS, Wastling JM. The proteome of *Toxoplasma gondii*: integration with the genome provides novel insights into gene expression and annotation. *Genome Biol.* 2008;9(7):R116.

Xu JH, Qin ZH, Liao YS, Xie MQ, Li AX, Cai JP. Characterization and expression of an actin-depolymerizing factor from *Eimeria tenella*. *Parasitol Res.* 2008 Jul;103(2):263-70.