Biological and Transcriptomic Comparison of Two Immunologically Distinct Strains of Eimeria maxima (GS and M6) and Characterization of Their Glycophosphatidylinositol (GPI) Anchored Surface Antigen Expression

by

Riadh Rahma Kazim Al-Badri

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Doctor of Philosophy
in
Pathobiology

Guelph, Ontario, Canada

© Riadh Rahma Kazim Al-Badri, August, 2013
ABSTRACT

BIOLOGICAL AND TRANSCRIPTOMIC COMPARISON OF TWO IMMUNOLOGICALLY DISTINCT STRAINS OF *EIMERIA MAXIMA* (GS AND M6) AND CHARACTERIZATION OF THEIR GLYCOPHOSPHATIDYLINOSITOL (GPI) ANCHORED SURFACE ANTIGEN EXPRESSION

Riadh Rahma Kazim Al-Badri

University of Guelph, 2013

Advisor: Dr. John R. Barta

Two immunologically distinct strains of poultry coccidium *Eimeria maxima*, Guelph (GS) and M6 strains, were investigated. Paired *in vivo* experiments demonstrated that GS and M6 have prepatent periods of approximately 120 h followed by peak oocyst shedding at 144-150 h post-inoculation. Fecundity of *E. maxima* M6 (12.8×10^3±1.95 oocysts shed/oocyst inoculated) was approximately twice that of GS (6.9×10^3±3.33) when inoculated with 1×10^3 infective oocysts per bird. Numerous sequential observations of synchronized populations of oocysts sporulating at 26°C showed no difference in the sporulation kinetics of the two strains; in both strains, sporogony was divided into five morphologically distinguishable stages whose abundance peaked at the following times during sporulation: unsporulated oocysts at 0 h; sporoblast anlagen at 18 h; sporoblasts without sporocyst walls at 22 h; and sporocysts without mature sporozoites at 38 h. Total RNA was isolated from four stages of sporogonic development (18 h, 22 h and 38 h of sporulation, and excysting sporozoites). These RNA samples were quantitatively pooled from each strain separately prior to selection of poly-A mRNA that was then fragmented, end-labeled and pyrosequenced using an Illumina HiSeq 2000 sequencer. The resulting transcriptome sequences (~48.8×10^9 bp total reads) were paired and *de novo* assembled. Ten thousands transcripts (5,000 from each strain) were searched against GenBank using blastx. A total of 2,067 transcripts of GS and 1,610 transcripts of M6 were assigned to putative biological function; ~60% of functionally annotated transcripts mapped to metabolic or cellular processes. GPI-anchored surface antigens (SAgs) identified in GS (18 SAgs) and M6 (18 SAgs) belonged to four major multi-copy gene families and 2 single-copy loci. Relative expression of SAgs expressed by both strains was generally similar; however, 3 GPI-anchored SAgs were uniquely expressed by each of GS and M6. One multigene locus demonstrated strain-specific SAg expression that may explain the lack of cross-immunity between these strains. This represents the first transcriptome data of sporulation of *E. maxima* and first comparison of immunologically distinct strains of any *Eimeria* sp. These data should aid in the search for antigenic targets that could be included in future subunit vaccines against these important agricultural parasites.
ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my advisor Dr. John Barta for the continuous support of my PhD study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I also would like to express my gratitude to my advisory committee members, Dr. Janet MacInnes and Dr. Lucy Mutharia, for their useful comments, remarks and engagement through the learning process of this PhD thesis.

Furthermore I also would like to thank Julie Cobeán for her technical help and support. Dr. Sergio Periera and Dr. Daniele Merico of the Center for Applied Genomics, Hospital for Sick Children for his guidance related to Illumina sequencing. Dr. Fiona Tomley (Royal Veterinary College, Hatfield, UK) and Dr. Adam Reid (Parasite Genomics, Wellcome Trust Sanger Institute, Cambridge, UK) are thanked for their willingness to share some of their unpublished work related to the nomenclature of *Eimeria maxima* genes. The research presented in this thesis was funded through various research grants awarded to my advisor from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Ontario Ministry of Agriculture and Food (OMAF). Presentation of some of my work at the American Society of Parasitologists’ Annual Meeting was supported through a Marc Dresden Travel Award from that Society.

I am thankful to the Iraqi ministry of higher education and scientific research for the generous scholarship they provided to me. I also would like to thank the Iraqi cultural attaché in Ottawa and, in particular, Dr. Assad Toma for his support and help.

Most importantly, I would like to thank my loved ones, colleagues, and friends who have supported me throughout entire process, both by keeping me harmonious and helping me putting the pieces together.
DECLARATION OF WORK PERFORMED

All work reported in this thesis was performed by myself and my advisor Dr. John Barta.
Table of Contents

TABLE OF CONTENTS ........................................................................................................V
LIST OF FIGURES ............................................................................................................. VII
LIST OF TABLES ............................................................................................................... VIII
LIST OF APPENDICES .................................................................................................... IX

CHAPTER 1 - LITERATURE REVIEW .............................................................................. 1
1.1 Coccidiosis and *Eimeria* ...................................................................................... 1
1.2 Life cycle and morphology of oocysts .................................................................. 2
  1.2.1 Life cycle ........................................................................................................ 2
  1.2.2 Morphology of the Oocyst............................................................................. 3
1.3 Vaccines against avian coccidiosis ..................................................................... 5
  1.3.1 Live unattenuated vaccines ......................................................................... 5
  1.3.2 Live attenuated vaccines .............................................................................. 6
  1.3.3 Recombinant protein vaccines ................................................................... 9
1.4 *Eimeria maxima* .................................................................................................. 14
  1.4.1 Coccidiosis caused by Eimeria maxima ..................................................... 14
  1.4.2 Immunogenicity and immunological diversity of Eimeria maxima .......... 15
1.5 Sporozoite as a target of immune responses ...................................................... 18
1.6 Thrombospondin related adhesive protein (TRAP) .......................................... 19
1.7 Analysis of genetic variation in *Eimeria* ............................................................ 20
  1.7.1 Multilocus enzyme electrophoresis (MEE) ............................................... 21
  1.7.2 Southern blot analysis ................................................................................ 22
  1.7.3 Analysis of chromosomal DNA .................................................................. 23
  1.7.4 One-dimension and 2-dimension SDS-polyacrylamide gel electrophoresis 24
  1.7.5 Amplified fragment length polymorphism (AFLP) .................................... 25
1.8 Transcriptome sequencing ....................................................................................... 26

CHAPTER 2 - INTRODUCTION AND HYPOTHESES .................................................... 28
2.1 Research Problem ................................................................................................ 28
2.2 Likely issues ......................................................................................................... 28
2.3 Hypotheses .......................................................................................................... 29
2.4 General approach .................................................................................................. 30

CHAPTER 3 - THE KINETICS OF OOCYST SHEDDING AND SPORULATION IN TWO IMMUNOLOGICALLY DISTINCT STRAINS OF *Eimeria maxima*, GS AND M6 .......... 31
3.1 Abstract ................................................................................................................ 31
3.2 Introduction .......................................................................................................... 31
3.3 Materials and methods ....................................................................................... 32
  3.3.1 Experimental animals and parasites ......................................................... 32
  3.3.2 Determining the peak of oocyst shedding ............................................... 33
  3.3.3 Kinetics of oocyst sporulation ................................................................. 34
3.4 Results .................................................................................................................. 35
  3.4.1 Comparison of oocyst shedding of E. maxima GS and M6 strains ........ 35
CHAPTER 4 - FUNCTIONAL ANNOTATION AND MAPPING OF TRANSCRIPT SEQUENCES OF TWO IMMUNOLOGICALLY DISTINCT STRAINS (GUELPH AND M6) OF THE CHICKEN COCCIDUM EIMERIA MAXIMA ................................................... 44
4.1 ABSTRACT ............................................................................................................. 44
4.2 INTRODUCTION .................................................................................................... 44
4.3 MATERIALS AND METHODS .............................................................................. 46
  4.3.1 Parasite Propagation ...................................................................................... 46
  4.3.2 Oocysts sporulation and bleaching ................................................................. 46
  4.3.3 RNA preparation ............................................................................................ 47
  4.3.3.1 Samples used for RNA extraction ................................................................. 47
  4.3.3.2 Extraction of RNA ....................................................................................... 47
  4.3.3.3 Further RNA purification ............................................................................. 47
  4.3.3.4 RNA quality assessment using Formaldehyde Gel Electrophoresis .......... 48
  4.3.4 Amplification of TRAP gene (PCR) ................................................................. 49
  4.3.5 cDNA Library Construction and Illumina (NextGen) Sequencing .......... 51
  4.3.6 De novo assembly and Blast2GO analysis ................................................... 52
4.4 RESULTS .............................................................................................................. 53
  4.4.1 RNA Isolation and Assessment of Quality ................................................... 53
  4.4.2 Illumina (NextGen) Sequencing ................................................................. 55
  4.4.3 Transcript Generation and Annotation from E. maxima GS and M6 .......... 55
4.5 DISCUSSION ....................................................................................................... 62

CHAPTER 5- IDENTIFICATION OF GLYCOPHOSPHATIDYLINOSITOL (GPI) ANCHORED SURFACE PROTEINS AND COMPARISON BETWEEN TWO IMMUNOLOGICALLY DISTINCT STRAINS OF EIMERIA MAXIMA (GS AND M6) ............................................................. 67
5.1 ABSTRACT ............................................................................................................. 67
5.2 INTRODUCTION .................................................................................................... 67
5.3 MATERIALS AND METHODS .............................................................................. 70
  5.3.1 Complementary DNA construction ............................................................... 70
  5.3.2 De novo assembly of paired reads ................................................................. 70
  5.3.3 Mapping of transcripts against sequences from E. maxima H and Wey ...... 71
  5.3.4 Relative abundance of gene (transcript) expression ..................................... 72
  5.3.5 Surface antigen (SAgs) characterization ..................................................... 73
5.4 RESULTS ............................................................................................................... 74
5.5 DISCUSSION ....................................................................................................... 91

CHAPTER 6 – GENERAL DISCUSSION ........................................................................ 96
CHAPTER 7 – CONCLUSIONS .................................................................................... 104
REFERENCES ............................................................................................................ 106
APPENDICES ............................................................................................................. 120
List of Figures

**FIGURE 1.1** Life cycle of a typical eimeriid parasite .......................................................... 3
**FIGURE 1.2** Oocyst structure .................................................................................................. 4
**FIGURE 1.3** Morphology of an *Eimeria maxima* oocyst ....................................................... 4
**FIGURE 3.1** Oocysts passed through 5 morphologically identifiable stages .......................... 35
**FIGURE 3.2** Pattern of oocyst shedding by *Eimeria maxima* GS and M6 ......................... 37
**FIGURE 3.3** Mean total oocyst output of *Eimeria maxima* GS and *E. maxima* M6 ........... 38
**FIGURE 3.4** Kinetics of oocyst sporulation at 26°C of *Eimeria maxima* GS and M6 .......... 39
**FIGURE 4.1** RNA integrity assessed using agarose formaldehyde gel electrophoresis ......... 54
**FIGURE 4.2** ‘TRAP’ (EmTFP250) amplified from M6 cDNA to test cDNA quality .......... 54
**FIGURE 4.3** Distribution of transcript lengths of *E. maxima* GS and M6 ............................. 56
**FIGURE 4.4** Characterization of transcripts generated from *E. maxima* GS and M6 ......... 57
**FIGURE 4.5** Biological process distribution of 2,067 transcripts from *E. maxima* GS ....... 59
**FIGURE 4.6** Biological process distribution of 1,610 transcripts from *E. maxima* M6 ...... 60
**FIGURE 4.7** Top hits species distribution of 10,000 transcripts of *E. maxima* M6 and GS .... 61
**FIGURE 5.1** Typical surface antigen gene, transcript and translation product .................... 75
**FIGURE 5.2** Multicopy Gene Family 1 - Type A ................................................................. 80
**FIGURE 5.3** Type A SAg transcripts of *Eimeria maxima* GS, M, H and Wey .................. 81
**FIGURE 5.4** Multi-copy gene family - Type B ....................................................................... 82
**FIGURE 5.5** Multi-copy gene family - Types C, D, J, and G ................................................ 83
**FIGURE 5.6** Multi-copy gene family - Type I ........................................................................ 84
**FIGURE 5.7** Single copy gene - Type E ................................................................................ 85
**FIGURE 5.8** Single copy gene - Type H ................................................................................ 85
**FIGURE 5.9** Aligned, annotated Type A surface antigen predicted translation products .... 87
**FIGURE 5.10** Abundance of gene expression of 36 GPI-anchored SAGs ............................. 89
**FIGURE 6.1** Biosynthesis of glycosylphosphatidylinositol (GPI)-anchored proteins ........... 98
List of Tables

TABLE 1.1 Commercially available live coccidiosis vaccines for chickens..........................6
TABLE 1.2 Commercially available and experimental attenuated live coccidiosis ..............8
TABLE 1.3 Subunit Vaccines .................................................................................................12
TABLE 1.4 Experimental DNA-based vaccines....................................................................14
TABLE 4.1 Total RNA extracted.............................................................................................62
TABLE 4.2 Percentages of transcripts for which putative biological function ....................58
TABLE 5.1 Description of 36 transcripts.................................................................................76
TABLE 5.2 Comparison of coding region .............................................................................90
List of Appendices

APPENDIX 3.1 Total oocyst counts of M6 and GS strains of *Eimeria maxima* ....................120

APPENDIX 3.2 One hundred viable oocyst counts every 4 hours of oocyst sporulation of
*Eimeria maxima* M6 ...........................................................................................................121

APPENDIX 3.3 One hundred viable oocyst counts every 4 hours of oocyst sporulation of
*Eimeria maxima* GS ...........................................................................................................122

APPENDIX 4.1 MOPS running buffer preparation ................................................................123

APPENDIX 4.2 Example of Blast2Go mapping of *Eimeria maxima* M6 transcripts ............124
Chapter 1 - Literature Review

1.1 Coccidiosis and *Eimeria*

Coccidiosis is a disease caused by obligate protozoan parasites belonging to the genus *Eimeria*. This disease infects the intestine of a host and reduces the growth efficiency. Coccidiosis is an important disease in poultry and other livestock (Lillevold and Trout 1993). Avian coccidiosis is caused by any of seven species of the genus *Eimeria* (*E. acervulina*, *E. brunetti*, *E. maxima*, *E. necatrix* and *E. tenella*, *E. mitis* and *E. praecox*). Coccidiosis is considered one of the most important diseases that affect the poultry industry worldwide. It is estimated that the annual costs of this disease are 800 million US$ as a result of reduced efficiency of feed conversion, high mortality rates, and the need for prophylactic drugs (Williams 1998).

Seven *Eimeria* species are important in chickens and can cause clinical disease. *Eimeria acervulina*, *E. brunetti*, *E. maxima*, *E. necatrix* and *E. tenella* produce moderate to severe intestinal lesions whereas *E. mitis* and *E. praecox* are infrequently pathogenic in the commercial poultry industry. There are two additional species of *Eimeria* (*E. hagani* and *E. mivati*) found in chickens but they are generally inapparent and rarely cause clinical coccidiosis. Moreover, the validity of *E. hagani* as a species has been questioned.

*Eimeria* species localize to different parts of the intestinal tract; for instance, *E. acervulina* usually localizes the duodenum and upper jejunum whereas *E. brunetti* infects the lower ileum and proximal areas of the caeca. *Eimeria maxima* and *E. necatrix* localize in the lower part of the duodenum to the ileum. *Eimeria necatrix* oocysts develop in the caeca which can also be infected by all stages of *E. tenella* (Tyckowski et al. 1991).

Clinical signs vary depending on the species of *Eimeria* but there are common signs for the disease in chicken. The infected birds tend to be depressed and huddle. As a result of decreased food and water consumption and decreased feed efficiency, birds tend to lose weight. The most important clinical sign with *E. necatrix* and *E. tenella* infection is bloody diarrhoea (Tyckowski et al. 1991).

The histopathological lesions differ between species of *Eimeria*. Clinically, diarrhoea, severe weight loss, poor feed conversion and loss of skin pigmentation are
common in *Eimeria* infected birds. During severe infection with *E. maxima*, the intestinal wall becomes thickened. The serosal surface of the intestine tends to be hemorrhagic. The intestine may be balloononed. Microscopically, numerous gametocytes and oocysts can be seen within the infected tissue (Tyckowski et al. 1991).

1.2 Life cycle and morphology of oocysts

1.2.1 Life cycle

*Eimeria* spp. have a complex life cycle divided into intracellular and extracellular stages with both sexual and asexual reproduction (Figure 1.1). The parasite typically develops through three phases: sporogony, merogony and gametogony (Hammond 1973). Sporogony occurs outside the host. In this phase, unsporulated oocysts develop to form sporulated oocysts. Sporulated oocysts have four sporocysts that each contains two sporozoites (Hammond 1973).

After ingestion of sporulated oocysts, sporocysts are released from the oocyst wall through the grinding action of the gizzard and the enzymatic action of the upper intestinal tract. Ultimately sporozoites are released (excyst) from sporocysts and become free in the intestine of chickens. After sporozoites invade the intestinal epithelium, they enter the next phase of the life cycle (merogony). During this phase, the sporozoite rounds up to form a trophozoite followed by nuclear divisions to form an immature meront (schizont). The cytoplasm of the meront divides to form a number of merozoites (Hammond 1973).

Gametogony begins when the merozoites enter cells and start transforming into microgamonts and macrogamonts. The microgamonts undergo multiple divisions producing numerous motile biflagellated microgametes (mature male gametes); whereby one macrogamonts gives rise to single macrogamete (mature female gamete). Motile microgametes invade cells and fertilize macrogametes found inside infected cells. Finally, wall forming bodies begin forming the oocyst wall (Hammond 1973).
Figure 1.1 Life cycle of a typical eimeriid parasite. Sporozoites (A) excyst from sporocysts in the digestive tract of the definitive host. Sporozoites penetrate host cells and initiate merogony (B). Merozoites penetrate additional host cells and undergo further rounds of merogonic replication (C-D). After one or more cycles, merozoites penetrate cells (E) and produce macrogametocytes (F) or microgametocytes (G). The latter locate and fertilize (H) the mature macrogamete. The resulting zygote, an unsporulated oocyst, is shed in the faeces (I). Sporogonic development occurs in the environment to produce a fully sporulated, infective oocyst (I-L). Reproduced with permission from Barta (2001).

1.2.2 Morphology of the Oocysts

The oocyst is the resistant (cyst) stage of the parasite. The mature sporulated oocyst wall is composed of one or two layers. In addition, the wall may be lined by a thin membrane. *Eimeria* oocysts have four sporocysts and each sporocyst contains two sporozoites. Oocysts may have a micropyle that may or may not be covered by a cap. Oocysts may or may not have refractile polar granules. Some material generally referred to as oocyst or sporocyst residua may be also observed. These residua form as a result of sporozoite and sporocyst formation. Sporocysts have Stieda bodies at their apexes. Sporozoites have small and large refractile globules. The structure of mature and sporulated *Eimeria* oocysts is shown in (Figure 1.2 and Figure 1.3).
Figure 1.2 Oocyst structure: Notice the spherical to sub-spherical shape with four sporocysts that contain two sporozoites in each one (from Duszynski and Wilbur 1997).

Figure 1.3 Morphology of an *Eimeria maxima* oocyst: Notice the spherical to sub-spherical shape of the oocyst. Within the oocyst, there are four sporocysts that contain two sporozoites in each. [Source: S. J. Upton, Kansas State University, Manhattan KS, with permission.]
1.3 Vaccines against avian coccidiosis

1.3.1 Live unattenuated vaccines

Two commercially live (virulent) anticoccidial vaccines, Immucox® (Vetech Laboratories, Inc., Buffalo, NY) and Coccivac® (Mallinckrodt Veterinary, Inc., Hillsboro, DE; Schering-Plough Animal Health) are currently available. These vaccines can induce long lasting protective immunity against *Eimeria* infections. The antigenic stimulation of the chick’s immune system is limited because of the inherent pathogenicity of these live vaccines. Moreover, it is known that host immunity is specific to a particular parasite species so live vaccines must contain numerous *Eimeria* species. This fact leads to several problems such as introducing new pathogenic *Eimeria* species into previously unexposed chickens, balancing infective dose with the pathogenicity of *Eimeria* species that are included in the vaccine, and the antigenic variability among *Eimeria* spp. strains present in the vaccines and strains present in the field (Martin et al. 1997).

The protective immunity produced by live vaccines depends upon live parasites that can cause infection in the host. Several cycles of fecal shedding and re-ingestion of oocysts follow vaccination lead to a slow increase in exposure of the birds to oocysts. Because vaccinated chickens are challenged and exposed to oocysts for long time, immunity is gradually developed over period of time. In this case, the parasite load and level of immunity are simultaneously increased (Martin et al. 1997).

Several methods of immunization have been studied to improve the efficacy of live vaccination such as using different vaccine delivery vehicles, changing the site of vaccine inoculation, introducing coccidiostats with vaccines, and vaccinating chickens with multiple low doses over a long period of time (trickle immunizations). Danforth et al. (1997) found that gel immunization technique is superior to immunization by gavage, by spray cabinet, or by the conventional delivery method of Immucox®.
<table>
<thead>
<tr>
<th>Abstract</th>
<th>Stage</th>
<th>Examples</th>
<th>Targeted Species</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delivery of a small number of virulent <em>Eimeria</em> sp. oocysts to young birds to elicit protection against challenge with the same species of parasite. This method requires ‘cycling’ of oocysts in vaccinate birds.</td>
<td>Whole virulent oocysts</td>
<td>IMMUCOX® FOR CHICKENS I</td>
<td><em>E. acervulina</em>, <em>E. maxima</em>, <em>E. necatrix</em>, and <em>E. tenella</em></td>
<td>Lee et al. 2009</td>
</tr>
<tr>
<td>One group of chickens were orally vaccinated with <em>Eimeria</em> species oocysts. Other groups were given anticoccidial drugs (Diclazuril and Toltrazuril). The vaccine performed better in protecting chickens from infection after they had been challenged.</td>
<td>Whole virulent oocyst (precocious line of <em>E. tenella</em>)</td>
<td>IMMUCOX® FOR CHICKENS II</td>
<td><em>E. acervulina</em>, <em>E. maxima</em>, <em>E. necatrix</em>, <em>E. tenella</em>, and <em>E. brunetti</em></td>
<td>Lee et al. 2009</td>
</tr>
<tr>
<td>Several <em>Eimeria</em> species oocysts were inoculated (through air cell) to ovo (18 days embryo). Sporozoites also inoculated to another group of ovo. The hatched chicks performed well when they were challenged. The vaccine reduced the oocyst output, reduced lesions, and improved weight gain.</td>
<td>Live oocyst or sporozoites</td>
<td>Coccivac®-D</td>
<td><em>E. acervulina</em>, <em>E. maxima</em>, <em>E. necatrix</em>, <em>E. hagani</em>, <em>E. praecox</em>, <em>E. tenella</em>, <em>E. brunetti</em>, and <em>E. mivati</em>.</td>
<td>Suo et al. 2006</td>
</tr>
</tbody>
</table>

### 1.3.2 Live attenuated vaccines

There are two commercially available live attenuated anticoccidial vaccines, Livacox® and Paracox®. Several studies have proved the safety and effectiveness of Paracox® in controlling coccidiosis (Waldenstedt et al. 1999). Attenuation of live parasites prior to inclusion in a vaccine is safer than using pathogenic field strains. Parasites can be
attenuated in several ways, such as selection for early development of oocysts (i.e., precocious development), parasite irradiation, and passage of parasites in embryonated eggs. Precocious strains are selected by repeatedly passing virulent parasites in vivo and collecting only the oocysts shed early in patency. After repeated passages, precocious strains of parasites are produced. These strains frequently produce fewer asexual merogonic generations which reduce the reproductive and pathogenic potentials of the parasite (Lillehoj and Lillehoj 2000).

Precocious strains are different from their parental strains. They display a reduced prepatent period, reduced reproductive potential, transformed intracellular development, and diminished infectivity (McDonald et al. 1986). McDonald and Shirley (1987) found that precocious strains of E. acervulina and E. tenella had smaller oocysts and generated fewer merozoites compared with their parental strains. A strain of E. necatrix selected for precocious development that exhibits decreased reproduction, prepatent period, and pathogenicity has also been described by Montes et al. (1997).

Embryonated chickens eggs have been used successfully to attenuate E. tenella, E. mitis, and E. necatrix, but not E. praecox, E. acervulina or E. maxima (Shirley and Long 1990). Although passing parasites in eggs does reduce virulence, this method of vaccine production has largely ceased. Another way of decreasing the pathogenicity of Eimeria relies upon adapting the parasites to grow and develop in cells grown in tissue culture. Zhang et al. (1997) adapted a field isolate of E. tenella to grow in primary chicken kidney cells and succeeded in producing 14 generations in vitro by repeated passages between cultured cells and chickens. In contrast with the original field strain, chickens infected with the cell-adapted parasite generally experience less pathogenic effects in terms of intestinal lesions, body weight, and mortality rates. An avian coccidiosis vaccine obtained from the supernatants of SB-CEV-1/F7 cells infected with E. tenella has been described by Brake et al. (1997). This vaccine confers a protection against challenge with the homologous parasite.

Irradiation is another method that has been used to attenuate oocysts (Gilbert et al. 1998). Augustine et al. (1993) found a significant improvement in weight gain and feed efficiency in chickens that were challenged with E. tenella after they had been inoculated
### Table 1.2 Commercially available and experimental attenuated live coccidiosis vaccines for chickens

<table>
<thead>
<tr>
<th>Abstract</th>
<th>Stage or Antigen</th>
<th>Examples</th>
<th>Targeted Species</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chickens were orally inoculated with live attenuated ionophore-resistant <em>Eimeria</em> oocysts. Chickens given anticoccidial drugs with the vaccine showed more resistance to infection than chickens supplied anticoccidial drug or vaccine alone.</td>
<td>Whole attenuated oocysts</td>
<td><em>E. tenella</em>, <em>E. maxima</em>, and <em>E. acervulina</em></td>
<td>Li et al. 2005</td>
<td></td>
</tr>
<tr>
<td>Vaccinated chickens were challenged with virulent <em>Eimeria</em> spp. (<em>E. acervulina</em>, <em>E. maxima</em>, <em>E. mitis</em>, and <em>E. tenella</em> or <em>E. necatrix</em>). The vaccinated chickens were protected showing no weight loss and lower intestinal lesions than unprotected chickens.</td>
<td>Whole attenuated oocysts</td>
<td><em>E. acervulina</em>, <em>E. maxima</em>, <em>E. mitis</em>, and <em>E. tenella</em>.</td>
<td>Williams 2003</td>
<td></td>
</tr>
<tr>
<td>One day old chicks were vaccinated with paracox-5 vaccine. They were challenged with the same <em>Eimeria</em> species but virulent. The immunized chicks showed less intestinal lesion and weight loss than unprotected chicks.</td>
<td>Whole attenuated oocysts</td>
<td><em>E. acervulina</em>, <em>E. maxima</em>, <em>E. mitis</em>, or <em>E. tenella</em></td>
<td>Crouch et al. 2003</td>
<td></td>
</tr>
<tr>
<td>Live attenuated vaccine was used in this study to evaluate its potency to protect chickens against <em>Eimeria</em> infection. The ability of the vaccine was also compared to using anticoccidial drugs such as (Narasin) or antibiotic (Virginiamycin) in protecting chickens. The vaccinated birds showed low oocysts production and satisfactory level of immunity. <em>Eimeria maxima</em> oocysts were attenuated by gamma-irradiation. Vaccinated chickens were challenged by virulent <em>E. maxima</em>. The chickens showed 10% increase in weight gain than unvaccinated chickens.</td>
<td>Whole attenuated oocysts</td>
<td><em>E. maxima</em></td>
<td>Jenkins et al. 1997</td>
<td></td>
</tr>
<tr>
<td>This vaccine was protecting only against homologous infection. Oocysts were attenuated by gamma radiation. This vaccine orally delivered to chickens subsequently challenged with virulent oocysts. The chickens showed decreased in oocyst production and were protected against further development of infection.</td>
<td>Whole attenuated oocysts</td>
<td><em>E. maxima</em></td>
<td>Shirley and Bellatti 1988</td>
<td></td>
</tr>
</tbody>
</table>

8
with irradiated *E. adenoeides*. Because chickens are foreign hosts for this turkey coccidium, *E. adenoeides* sporozoites excyst and invade the intestinal epithelium but undergo little further growth. X-irradiated oocysts of *E. acervulina* were used by Jenkins et al. (1991a) to induce protective immunity to a virulent challenge with the homologous parasite. They found that host protective immunity occurred after vaccination (during a critical early period of intracellular parasite development). Jenkins et al. (1997) found that chickens vaccinated with X-irradiated oocysts of *E. maxima* showed decreased development of asexual parasite stages in the host intestinal mucosa.

### 1.3.3 Recombinant protein vaccines

A critical step in recombinant vaccine development is identification of the life cycle stages that induce protective immunity. Because they are fairly easy to obtain and impeding their activities should theoretically prevent infection, sporozoites are the preferred parasitic form to target for development of recombinant protein vaccine. Brothers et al. (1988) illustrated an antigenic surface protein of *E. tenella* sporozoites using neutralizing monoclonal antibodies Ptn 7.2A4/4 and Ptn 9.9D12. The corresponding gene was cloned and expressed in *Escherichia coli*, but the immunogenicity of this protein was not studied. Miller et al. (1998) reported that the *E. acervulina* sporozoite recombinant antigen (GX3262) stimulated partial protective immunity after a single immunization in 2-day-old broiler chickens, as indicated by decreased cecal lesions in vaccinated chickens compared with similarly challenged non-immunized controls.

Several recombinant *Eimeria* proteins, e.g., 5401 (Danforth et al. 1989), MA16 (Castle et al. 1991) and GX3264 (Bhogal et al. 1992), have been used to produce candidate vaccines. It was found that these antigens stimulate partial resistance to subsequent challenge with live parasites. Cell surface antigens can be significant components of vaccines because of their function in host cell invasion (see Taberes et al. 2004; Jung et al. 2004; Gilson et al. 2006). Jenkins et al. (1998) have cloned and expressed a cDNA encoding an immunogenic region of a 22-kD surface protein of *E. acervulina* sporozoites. Lillehoj et al. (1988) found that the recombinant protein MA1 stimulated significant *in vitro* activation of T lymphocytes obtained from chickens inoculated with *E. acervulina*. 
An immunogenic region of a surface antigen shared between sporozoites and merozoites is encoded by the M16 protein of *E. acervulina* (Castle et al. 1991). T lymphocytes from *E. acervulina* immune chickens are activated *in vitro* by M16 expressed in *E. coli*. Jenkins et al. (1988) cloned genes encoding sporozoite (cSZ-1) and merozoite (cMZ-8) surface antigens of *E. acervulina* cDNA expression libraries. They also found that humoral and cellular immune responses *in vitro* were stimulated by the gene products of those genes. Likewise, Jenkins et al. (1990) found that humoral and cellular antigen responses were induced after oral immunization of chickens with a live *E. coli* transformant expressing the recombinant cMZ-8 protein. Parenteral inoculation of chickens with recombinant cMZ-8 protein alone was more effective in some strains of chickens. Inoculation of chickens with *E. coli* cMZ-8 transformants induced more effective local intestinal immunity in other strains.

Immunization of chickens with live *E. coli* transformants expressing recombinant proteins is a valuable way of stimulating protective immunity. Lillehoj et al. (1990) found that oral vaccination with live *E. coli* expressing the p250 surface antigen of *E. acervulina* or intramuscular inoculation with purified recombinant p250 causes intestinal and systemic antigen-specific T-cell and humoral responses *in vitro* and caused a decrease in mucosal parasitism compared with inoculation of *E. coli* alone. However, Kim et al. (1989) found that immunization of chickens with *E. coli* p250 transformants stimulated partial protective immunity against *in vivo* challenge with live coccidia. Jenkins et al. (1991b) reported that oral inoculation of live bacteria expressing the gene encoding an *E. acervulina* merozoite antigen (EAMZ250) to 1-wk-old chickens induced antigen-specific humoral and cellular immune responses. Moreover, important immune responses to *in vivo* challenge were obtained based on decreased weight loss and intestinal lesions.

The parasite antigens, proposed to be candidates for vaccines such as components of refractile bodies and apical organelles, have important roles in the parasite’s life cycle. Inhibiting these antigens by inducing immunity against them should theoretically inhibit parasite infectivity (Crane et al. 1991). Miller et al. (1989) found that *in vitro* development of schizonts was inactivated by a monoclonal antibody when their refractile body protein (p28) was countered. The possibility of developing a vaccine against coccidiosis was
shown with another recombinant refractile protein synthesized from *E. tenella* and designated as CheY-SO7. Two-day-old chicks were protected against coccidiosis caused by *E. tenella*, *E. acervulina*, or *E. necatrix* after they had been inoculated with a single low dose of recombinant antigen (CheY-SO7) in the absence of adjuvant. Cell mediated immune mechanisms elicited in immunized chickens protected chickens against subsequent challenge (Crane et al. 1991). This same antigen was able to protect chickens against an *E. tenella* challenge when introduced as a plasmid-mediated DNA vaccine construct (Kopko et al. 2000).

There are several organelles located at the anterior tip of coccidia (e.g., the rhoptry organelles) that play crucial role for the invasion and parasitophorous vacuole formation and survival of apicomplexan parasites such as *Eimeria* within host cells (see Tomley 1994; Entzeroth et al. 1997; Oakes et al. 2013). These organelles have been isolated from sporozoites and have a complex polypeptide profile with at least 60 separate proteins (Kawazoe et al. 1992). Tomley et al. (1996) used monoclonal antibodies that distinguished epitopes found in surface and rhoptry of *E. acervulina* to isolate recombinant proteins that stimulate T cells *in vitro*. Jenkins et al. (1990) synthesized cDNA encoding a merozoite specific *E. acervulina* antigen (EAMZp30-47). This antigen contains epitopes shared by several surface and rhoptry proteins. T cells in *E. acervulina* immune chickens were induced after the purified EAMZp30-47 had been inoculated.

Protein based vaccines produced to date are unable to induce completely protective immunity against challenges with field parasites. On the other hand, commercially available live vaccines such as Immucox® (Vetech Laboratories, Inc., Buffalo, NY) and Coccivac® (Mallinckrodt Veterinary, Inc., Hillsboro, DE; Schering-Plough Animal Health) have been shown to protect chickens against infection with *Eimeria*. These vaccines confer a high level of protection against subsequent infections with *Eimeria*. In addition, two commercially available live attenuated anticoccidial vaccines (Livacox® and Paracox®) have been shown to be efficacious. Several studies have proved the safety and effectiveness of Paracox® in controlling coccidiosis (Waldenstedt et al. 1999). Finally, live and attenuated coccidia vaccines are commercially available and have proved efficiency in
protecting of chickens against subsequent infections with various *Eimeria* species, more than commercially unavailable protein based vaccines.

Table 1.3 Subunit Vaccines (Protein or Vectored, e.g. *Salmonella*-based) and related studies to find a subunit vaccine against coccidiosis in chickens

<table>
<thead>
<tr>
<th>Study Summary</th>
<th>Stage or Antigen</th>
<th>Examples</th>
<th>Targeted Species</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purified gametocyte antigens were injected intramuscularly to breeding hens. Maternal antibodies (IgG) transferred from hens to their offspring reducing the oocysts production in the infected progenies</td>
<td>Purified gametocyte antigens (primarily EmTFP250)</td>
<td>E. maxima</td>
<td>Wallach et al. 2008</td>
<td></td>
</tr>
<tr>
<td>Microneme recombinant gene was inoculated to ova. The gene induced protective immunity against challenge with <em>E. tenella</em>. The titre of antibodies increased at 10-17 days following the challenge</td>
<td>Microneme recombinant gene (EtMIC2) and encoded protein</td>
<td>E. tenella</td>
<td>Ding et al. 2005</td>
<td></td>
</tr>
<tr>
<td>Gam56- gam82 genes were cloned in bacterial expression vector (pTRCHisB). The expressed and purified proteins were intramuscularly injected to chickens. High antibody response recognised by ELISA was achieved by using this recombinant protein</td>
<td>Purified gametocyte antigens (gam56-gam82)</td>
<td>CoxAbic</td>
<td>E. maxima</td>
<td>Belli et al. 2004</td>
</tr>
<tr>
<td>This protein was studied and corresponding cDNA was cloned and shown to encode lactate dehydrogenase enzyme. This protein was partially protective against <em>E. acervulina</em> infection. This protein was isolated from oocysts, sporozoites, schizonts, and merozoites but RNA was only isolated in the schizont stage</td>
<td>Purified protein (37 kDa)</td>
<td>E. acevulina</td>
<td>Schaap et al. 2004</td>
<td></td>
</tr>
<tr>
<td>This protein is member of TRAP family that is important in gliding motility and cell invasion by Apicomplexa members. <em>Escherichia coli</em> strains DH5-a and TOP10 were used for the propagation of recombinant expression plasmid. The immunogenicity of this protein was tested when it was inoculated to mice and chickens. It induces strong IgG responses in both models</td>
<td>Purified microneme protein (EmTFP250)</td>
<td>E. maxima</td>
<td>Witcombe et al. 2004</td>
<td></td>
</tr>
<tr>
<td>Chickens developed cell mediated immunity against <em>E. acervulina</em> infection after they had been vaccinated with recombinant 3-1E protein or direct injection of 3-1E cDNA</td>
<td>3-1 cDNA or its polypeptide (60kD and 23kD protein)</td>
<td>E. acervulina</td>
<td>Lillehoj et al. 2000</td>
<td></td>
</tr>
<tr>
<td>Study Summary</td>
<td>Stage or Antigen</td>
<td>Examples</td>
<td>Targeted Species</td>
<td>References</td>
</tr>
<tr>
<td>------------------------------------------------------------------------------</td>
<td>------------------------</td>
<td>-----------------------------------------------</td>
<td>------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Chickens were subcutaneously inoculated recombinant antigen that was isolated from <em>E. maxima</em> sporozoites and expressed in <em>Escherichia coli</em>. It was found that this recombinant antigen induce both antibody and lymphoproliferative responses</td>
<td>70 kDa protein</td>
<td><em>E. maxima</em></td>
<td>Bumstead et al. 1995</td>
<td></td>
</tr>
<tr>
<td>Sequence encoding 100 kDa refractile body protein was cloned in herpesvirus of turkeys (HVT). The protein was isolated from <em>E. acervulina</em>. Vaccinated chickens were challenged with <em>E. acervulina</em> and <em>E. maxima</em>. Significant improvement in weight gain was noticed in chickens vaccinated with recombinant HVT</td>
<td>100 kDa refractile body protein (Ea1A)</td>
<td><em>E. acervulina</em> and <em>E. maxima</em></td>
<td>Croneberg et al. 1999</td>
<td></td>
</tr>
<tr>
<td>Gametocyte antigen were prepared and used as a vaccine. Breeding chickens were intramuscularly inoculated the prepared antigen. Significant maternal IgG was produced. Progeny was protected against infection with <em>E. maxima</em>, <em>E. tenella</em>, and <em>E. acervulina</em> (based on oocysts output). This vaccine offered cross-protection</td>
<td>Gametocyte antigens</td>
<td><em>E. maxima</em></td>
<td>Wallach et al. 1995</td>
<td></td>
</tr>
<tr>
<td>The cross protection between several <em>Eimeria</em> species were studied in this trial. A recombinant antigen inoculated to chickens that were challenged subsequently by several <em>Eimeria</em> species. The vaccine offered protection against not only <em>E. tenella</em> but also against <em>E. maxima</em>, <em>E. acervulina</em>, and <em>E. necatrix</em></td>
<td>Recombinant antigen (CheY-SO7)</td>
<td><em>E. tenella</em></td>
<td>Grane et al. 1991</td>
<td></td>
</tr>
</tbody>
</table>
### Table 1.4 Experimental DNA-based vaccines against coccidiosis in chickens

<table>
<thead>
<tr>
<th>Abstract</th>
<th>Stage or Antigen</th>
<th>Target Species</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recombinant plasmid injected (intramuscular leg injection) to chickens to study the cross protection between several <em>Eimeria</em> species. Immune responses were stimulated alleviating the infection as a result of the cross reactivity between <em>Eimeria</em> species used in this trial (conserved antigens). The immune response was differentially protective among <em>Eimeria</em> species used in this trial.</td>
<td>Recombinant plasmid-DNA vaccine (pVAX1-cSZ2-IL-2)</td>
<td><em>E. tenella</em>, <em>E. necatrix</em> and <em>E. maxima</em></td>
<td>Shah et al. 2010</td>
</tr>
<tr>
<td>Chickens were intramuscularly injected DNA vaccine (antigen). The vaccine is protective against the challenge with <em>E. acervulina</em> and <em>E. necatrix</em> but not with <em>E. maxima</em>.</td>
<td>DNA (pcDNA-TA4-IL-2)</td>
<td><em>E. tenella</em>, <em>E. necatrix</em> and <em>E. acervulina</em></td>
<td>Song et al. 2009</td>
</tr>
<tr>
<td>Refractile body antigen was isolated from sporozoite of <em>E. tenella</em> and it was injected to chickens. The chickens challenged by virulent parasite. Decrease in cecal lesions and weight loss has been noticed in immunized chickens proving the efficiency of the antigen in inducing protective immunity.</td>
<td>Refractile body antigen</td>
<td><em>E. tenella</em></td>
<td>Kopko et al. 2000</td>
</tr>
</tbody>
</table>

### 1.4 *Eimeria maxima*

#### 1.4.1 Coccidiosis caused by *Eimeria maxima*

*Eimeria maxima* causes coccidiosis primarily in chickens affecting the mid-intestine. Gross lesions such as thickening of the intestinal wall, pinkish exudate that can fill the intestine, and obvious petechial lesions and haemorrhage are seen. The infection can extend from the duodenum to the intercaecal junction (Long 1973). Mild to severe diarrhoea and dehydration result from this haemorrhagic enteritis. Birds may lose body weight after four to six days post infection (Long 1959). The disease caused by *E. maxima* is considered moderate in comparison to the disease that caused by *E. tenella* or *E. necatrix*; however, the mortality rate can reach 20% when a chicken flock is severely affected by *E. maxima*.

*Eimeria maxima* was described before other coccidial parasites of chickens were studied (Tyzzer 1929). It was named *E. maxima* to reflect the large size of its oocysts,
which measure about 30 µm × 20 µm. This is the largest of the oocysts of *Eimeria* species found in chickens. The diagnosis of *E. maxima* may overlap with *E. brunetti* infection because they infect the mid-intestine and these two parasites cause similar pathological changes (Levine 1942). Epidemiological studies have demonstrated that *E. maxima* causes a disease in chickens worldwide based on using a variety of diagnostic methods such as studying sizes of oocysts, intestinal gross lesions, immunological cross-protection experiments and isoenzyme analysis. *Eimeria maxima* is one of the most frequently diagnosed coccidia in poultry worldwide (Williams et al. 1996).

Frequently mixed infection between *Eimeria* spp. is common among chickens. Although not studied specifically, co-infection with other *Eimeria* species is likely. Live vaccination is one case of intentionally multi-species infections and acquisition of immunity is not affected as far as is known although this phenomena has not been studied systematically.

### 1.4.2 Immunogenicity and immunological diversity of *Eimeria maxima*

A study by Rose and Long (1962) to investigate the immunogenicity of *E. maxima* showed that it is the most immunogenic species of avian coccidia. This observation was confirmed later by other researchers (e.g. Joyner and Norton 1976; Long and Millard 1977). In fact, partial immunity against *E. maxima* infection in chickens was achieved by inoculation with only a single sporocyst of *E. maxima* (see Lee and Fernando 1978).

Infection with *Eimeria* species induces strong protective immune responses that usually protect a particular host from subsequent infection with the same *Eimeria* species. Even though a single infective oocyst of *E. maxima* can induce strong immune responses that protect chickens against subsequent infections, the level of protection to the subsequent challenges still depends on the frequency and the level of the initial exposure to parasite. Administering a single *E. maxima* oocysts, numerous times can induce protective immunity and eliminate the subsequent infection more effectively than inoculating the same number of the oocysts in a single dose at one time. In a study by Joyner and Norton (1976), 1-5 oocysts of *E. maxima* or 5-20 *E. acervulina* were inoculated daily into chickens; and the inoculation of oocysts was ceased after oocyst production stopped.
Subsequently, the chickens were challenged with both species. The level of immunity with serial inoculations was stronger and/or lasted longer than that produced by a single inoculation of oocysts.

Many studies have examined the cross-protection amongst the seven *Eimeria* species infecting chickens. Infection with *E. maxima* does not induce protective immunity in the host against infection by any other *Eimeria* species. However, some studies have suggested that there is a level of immunological cross-protection between *E. maxima* and *E. brunetti* based on oocyst production (Rose 1967). There are some similarities between these species in terms of the degree of immunogenicity and the shape and size of their oocysts. Phylogenetic reconstructions using complete small subunit ribosomal DNA sequences (Barta et al. 1997), and more recently using partial sequences of the mitochondrial cytochrome oxidase I gene (Ogedengbe and Barta 2012), have demonstrated that *E. maxima* and *E. brunetti* are closely related but distinct species infecting the same host. As far as is known, these parasites are reproductively isolated.

Strain-specific protective immunity has been identified in some *E. maxima* strains by inoculating a single representative strain/isolate into birds that were later challenged with several strains/isolates of *E. maxima* (Norton and Hein 1976; Adrian et al. 2002). Although some birds are completely immune against infection with homologous strains/isolates following vaccination, these same birds may produce millions of oocysts when they are challenged with immunologically heterologous strains/isolates of the same species. For example, the cross-protection between the Houghton and Weybridge strains isolated in the UK and an *E. maxima* strain isolated in Malaysia was between 50 and 90% (Long 1974). Martin et al. (1997) obtained similar results as demonstrated by lesion score birds at post-mortem with a number of North American *E. maxima* strains. Similar results measuring oocyst output following homologous and heterologous in vivo challenge was obtained by Norton and Hein (1976) when they studied the cross-protection among various *E. maxima* strains.

It is well known that *E. maxima* strains are immunologically variant and therefore live vaccines may fail to protect chickens against infections with heterologous strains. In other words, the parasites that might be used in a live vaccine may fail to protect chickens
against infection with immunologically unrelated parasite populations even though they belong to the same *Eimeria* species, *E. maxima*. Parasites used in a live vaccine called Immucox® did not succeed in protecting Immucox®-immunized chickens when they were challenged by an immunologically unrelated *E. maxima* strain. It was thought that Immucox® was unable to protect chickens against the variant strain but the problem was resolved simply by adding a small number of oocysts of the immunologically variant strain of *E. maxima* to the vaccine (Immucox®) (Schnitzler and Shirley 1999). Although *Eimeria* spp. have an obligate sexual life cycle and therefore genetic recombination can and does occur (e.g. Jeffers 1976), single haploid sporozoites are capable of initiating an infection and producing viable oocysts in vivo in chickens (Shirley and Millard 1976). These observations suggest that *Eimeria* species of chickens (and likely other monoxenous coccidia) can both generate novel genotypes of parasites through genetic recombination during sexual reproduction as well as rapid clonal expansion of these novel genotypes when a mixed population is put under selective pressure (e.g. by anticoccidial usage or live vaccine usage).

The problem of immune variation of *E. maxima* has been addressed by using more than one representative *E. maxima* strain in the vaccine. For instance, two immunologically distinct strains of *E. maxima* have been included in a live vaccine called Paracox® (Shirley and Bellatti 1988). Long and Millard (1979) have demonstrated that immunizing chickens with a mixture of variant strains of *E. maxima* can stimulate efficient protective immunity against infection with several variant strains of *E. maxima*. Paracox® can confer a wider range of protection than vaccines that include a single strain alone.

*Eimeria maxima* is the choice for experiments of an immunological nature (Adrian et al. 2002). Even small numbers of this parasite can induce complete protective immunity against re-infection with the same parasite species, particularly when they are homologous strains. The early asexual stages of *E. maxima* (and other species of *Eimeria*) up to second generation schizonts are the most important stages to induce protective immunity against natural coccidian infections in chickens. Jenkins et al. (1993) found that chickens were completely immune against infection after they had been orally inoculated with gamma irradiated oocysts. With these irradiated oocytes, the developing sporozoites invaded the
intestine but didn’t undergo schizogony (asexual reproduction by simultaneous multiple fission). Chickens immunized by irradiated sporozoites had sufficient immunity to prevent further development of sporozoites to complete the life cycle inside the host.

1.5 Sporozoite as a target of immune responses

Species-specific protective immune responses can be induced in the host following infection with all species of *Eimeria* that are immunologically variable. However, some species can only induce protective immunity after several infections. In contrast, *E. maxima* is considered highly immunogenic because ingestion of only a few oocysts can stimulate near complete protective immunity against homologous challenges. That said, little to no information is known about the parasite molecules that are responsible for induction of the immune response or which stage(s) of the coccidian life cycle drives protective response. Even though the sporozoite is a known target for attack in the immune host, the precise role of this stage in inducing protective immune responses is ambiguous (Rose et al. 1984).

Jeffers and Long (1985) found that sporozoites of *E. tenella* inhibited by the anticoccidial drug, decoquinate, stimulated significant protective immunity in chickens that were subsequently challenged with decoquinate resistant *E. tenella* sporozoites. The immunity induced eliminated the decrease in weight gain but provided only slight protection against intestinal lesions. Jenkins et al. (1991a) found that protective immune responses were elicited only after *E. tenella* sporozoite invasion. In this trial, chickens were inoculated with *E. acervulina* or *E. tenella* oocysts that had been gamma irradiated. It was found that the oocyst output and the presence of schizonts in the intestine were decreased during the infection due to the restriction in development of the early parasite stages.

Rose and Hesketh (1976) found that sexual stages are weakly immunogenic whereas the early endogenous stages are strongly immunogenic. In other words, they found that protective immunity was produced against the asexual stages of *Eimeria*, such as sporozoites, but incomplete protection was produced against the sexual stages of the parasite. It was concluded that the immunity against sexual stages probably does not play
an important role in limiting the onset of the disease but may have role in decreasing oocyst shedding.

Several studies have described the developmental stages of *Eimeria* that are affected by the immune system. For example, Leatham and Burns (1967) found significantly fewer *E. tenella* sporozoites in the lamina propria 12 h post-challenge in immune chickens compared to naive controls. In further studies, they observed no difference in the invasiveness of sporozoites between non-immune and immune chickens suggesting that there could be intestinal immune mechanisms eliminating sporozoite progression (migration and/or development) early in the invasion of the intestinal epithelium.

The early stages of the *Eimeria* life cycle are affected by the adaptive immune response in previously exposed chickens. Rose et al. (1984) found that the effect of immune response is more significant during the later stage of invasion when sporozoites of *E. tenella* transfer from the intra-epithelial lymphocytes (IEL’s) to crypt epithelial cells. In contrast, Riley and Fernando (1988) observed that sporozoites of *E. maxima* appear to be affected early after invasion of the intestinal epithelium. In the latter study, the effects of immunity on the sporozoites were observed within 24 h post challenge indicating quick identification and response to the early stages of the life cycle (Riley and Fernando 1988); these observations were later confirmed by Beattie et al (2001).

**1.6 Thrombospondin related adhesive protein (TRAP)**

Thrombospondin related adhesive protein (TRAP) is a well characterized type I-transmembrane protein that contains a thrombospondin type I repeat (TSR) domain, a von Willebrand factor (vWF)-like A domain, and a short acidic cytoplasmic tail domain (CTD). TRAP was first isolated from *Plasmodium falciparum* and homologs were later isolated from other *Plasmodium* species (Templeton and Kaslow 1997) and *E. maxima* (Witcombe et al. 2003). TRAP is a microneme-based protein that is important in motility and cell invasion in members of the phylum Apicomplexa. TRAP is secreted from the anterior tip of sporozoites during movement and spreads along the body surface of the sporozoites to reach the posterior part of the parasite during sporozoite movement or the host cell invasion process. TRAP plays an essential role in host cell invasion by binding to
the actin-myosin motor of the parasite via its cytoplasmic portion while binding to the host cell through its extracellular domain (Menard 2000).

Gene targeting studies have demonstrated the importance of TRAP protein in cell invasion and gliding motility. These studies have demonstrated that disrupting TRAP impairs gliding motility, host cell invasion and sporozoite infectivity (Sultan et al. 1997). This protein has been studied as a target that could induce protective immunity against different apicomplexan parasites such as Eimeria, Toxoplasma and Cryptosporidium. It was observed that TRAP can be inhibited by antibodies leading to disruption of cell invasion and reduced motility of the parasite, specifically the sporozoite (Spaccapelo et al. 1997). Thus, TRAP has been considered as a component of multivalent subunit vaccines (e.g. Moorthy et al. 2004).

1.7 Analysis of genetic variation in Eimeria

There is little information about the genetic and molecular variation among Eimeria species and strains infecting chickens and few studies have been done regarding genetic variation in Eimeria. The identification of species and strains of Eimeria is based on the morphology of oocysts or pathology and lesions caused by Eimeria in the chicken’s intestine. The need for more accurate techniques to study the genetic variation among Eimeria populations has been noted (see Schwarz et al. 2009).

The seven Eimeria species infecting chickens were differentiated from each other based on analysis of the mobility of metabolic enzymes through starch gel or isoelectric focusing (Smith et al. 1994). The application of enzyme electrophoresis such as multilocus enzyme electrophoresis (MEE) revealed restricted intra-specific polymorphism traits by which at least 70 E. tenella isolates can be distinguished (Shirley et al. 1989). In a recent study, Eimeria proteomes were compared by using one-dimensional and two-dimensional gel electrophoresis and little intra-specific variations among strain and larger inter-specific variations among species were observed (Sutton et al. 1989; Barta et al. 1998).
1.7.1 Multilocus enzyme electrophoresis (MEE)

Multilocus enzyme electrophoresis is based on separation of non-denatured proteins (e.g., enzymes) while they migrate through a gel under the effect of an electric field (Andrews and Chilton 1999). Because there is variation in the transcribed DNA sequences, homologous enzymes may be different in size and charge. Since the rate of a protein’s migration through a gel under the influence of electric fields varies in size and charge, enzymes collected from different species and individuals may be separated from each other while these enzymes are passing through the gel. Enzyme-specific staining can be used to expose the positions of the enzymes separated on the gel (Andrews and Chilton 1999).

MEE has been successfully used to study the antigenic variation within species and strains of *Eimeria* that infect chickens. Eleven strains of 6 species of *Eimeria* that infect chickens were investigated to detect the variation using MEE (Shirley 1975). It was found that four enzymes can be used to identify and differentiate each *Eimeria* species. Shirley also found some intra-specific variation between species of laboratory isolates of *Eimeria* species but MEE has also been shown to be valuable for the investigation of mixed field samples. The electrophoretic mobility of glucose phosphate isomerase (GPI) in reference strains of *E. acervulina*, *E. praecox* and *E. mivati* was compared to the electrophoretic banding profiles of *Eimeria* taken from litter from chicken farms. The tested species and strains have been effectively differentiated based on their profiles (Chapman 1982).

MEE has been successfully used to genetically differentiate *Eimeria* species infecting chickens in numerous laboratories around the world, such as Australia (Andrews et al., 1990), Canada and USA (Johnston and Fernando 1997), Japan (Nakamura et al. 1986) and Sweden (Thebo et al. 1998). Coccidia are well suited to this technique because of the large number of *Eimeria* oocysts that can be collected and pooled from the tested fecal samples. Although sometimes the number of *Eimeria* oocyst found in the feces is low, the initial oocysts sample can be amplified easily by passage through chickens (Long and Joyner 1984).

However, getting large numbers of oocysts is not always guaranteed and passage through chickens is relatively expensive and time-consuming (Fernandez et al. 2003a). This is a lengthy process that introduces long lead times before study results become
available (Procunier et al. 1993; Fernandez et al. 2003b). There is strong selection pressure for stability in protein function resulting in relatively limited variation in a particular protein's amino acid sequence. As MEE detects such variation, its application is restricted because of the limited magnitude of variation present (Fernandez et al. 2003a,b). Moreover, the amino acid sequence of a particular protein may not vary further, even though there may be significant variation in the underlying DNA sequences; even if the amino acid sequence of the protein does change, the change may not affect the mobility of the protein during gel electrophoresis (Shirley 1994a).

1.7.2 Southern blot analysis

This technique was developed in 1975 (Southern 1992). The principle of this technique is detection of specific fragments of DNA separated by electrophoresis. Restriction endonucleases are used to digest the genomic DNA to convert it to fragments. The resultant fragments, are hybridized to specific probes and then gel electrophoresed. After this, DNA is transferred to a membrane such as nitrocellulose fluorography or autoradiography. The southern blot reproducibility depends on the ability of restriction enzymes to cleave DNA at specific recognition sites in a nucleotide sequence.

This technique was first used to study *Eimeria* spp. by Ellis and Bumstead (1990). Three *Eimeria* species (*E. tenella, E. acervulina* or *E. necatrix*) were studied in this trial. The genomic DNA of these species were fragmented by using EcoRI and probed by RNA probes prepared from rRNA obtained from sporulated oocysts. The three species of *Eimeria* were easily differentiated from one another. However, Ellis and Bumstead (1990) also found that using other endonucleases, such as *BamHI* and a pRibI probe, cannot differentiate between *E. tenella* strains (Houghton, Weybridge, and Doran).

In a later study, Shirley (1994a) examined seven genetically different *E. tenella* strains using a number of endonucleases including *HinfI, Rsal, HpaII* and *EcoRI*. Four probes were generated from repetitive sequences of genomic DNA of *E. tenella* and used to hybridize to the fragments formed by separate digestions with the endonucleases. Shirley found variation between *E. tenella* strains Houghton and Wey that was not detected by Ellis and Bumstead (1990) using *BamHI* and a pRibI probe. The fact that intra-specific
variation was detected by using Southern blotting made this technique more efficient and effective than other techniques such as MEE, in studying the variation among *Eimeria* species and strains. The drawbacks from using this technique are that a probe to a known target must be generated, designed, and optimized to hybridize specifically to the DNA fragments. This technique is considered complex because many steps are involved and there is need for a large number of oocysts to extract large amounts of DNA (Procunier et al. 1993; Shirley 1994a; Fernandez et al. 2003b).

### 1.7.3 Analysis of chromosomal DNA

Both pulsed field gel electrophoresis (PFGE) and field inversion gel electrophoresis (FIGE) use electric current applied through a low density agarose gel to promote the migration and separation of large DNA molecules based on their sizes. Traditional gel electrophoresis can be used to separate large DNA molecules (up to 15 to 20 kb). However, by making some modifications in the traditional gel electrophoresis such as applying the electrical current as pulses, pulsed field gel electrophoresis has been developed to separate DNA molecules of 50 kb to several Mb. The positions of DNA molecules in the gel can be detected by hybridization to radioactive probe or using stains such as ethidium bromide (Fernando and Pasternak 1991; Shirley 1994b).

Shirley et al. (1990) studied the karyotype of *E. tenella* using PFGE. Although it had been proposed in previous studies that the *E. tenella* genome had only 5 chromosomes, at least 12 chromosomes were detected by using PFGE. In a subsequent study (Shirley 1994b); differences in the karyotypes of a number of *E. tenella* strains were detected with chromosomes demonstrating significant size differences between strains. Further, specific probes that bound to a particular chromosome in one strain did not always bind to the homologous chromosome in the other strains revealing that differences also existed between chromosomal homologs.

Field inversion gel electrophoresis (FIGE) is a technique has been used for the same purpose. Unlike PFGE which uses complicated patterns of diagonal and horizontal electric fields, FIGE simply inverts the polarity of the electrical field from forward to reverse and back again to assist the migration of large DNA molecules through the gel (Carle et al.
When FIGE was applied to study the molecular karyotype of five *Eimeria* species isolated from chickens, it was found that each species has a distinctive pattern (Fernando and Pasternak 1991).

Electrophoretic karyotyping techniques can offer tremendous information about the genome such as the minimum number and the size of nuclear chromosomes, which helps to compare *Eimeria* species and strains. Chromosomes separated by PFGE can bind to probes differentially providing additional information about the chromosomes to study and compare species and strains of *Eimeria* infecting chickens (Shirley 1994b). However, using these procedures to study the chromosomes cannot provide particularized and detailed information about the genome. In addition, these techniques are relatively time consuming and a single electrophoretic separation can take more than 10 days. This technique does need a large amount of DNA for analysis (Shirley 1994b).

### 1.7.4 One-dimension and 2-dimension SDS-polyacrylamide gel electrophoresis

These electrophoretic techniques separate charged molecules based on their physical properties such as charge or mass as they are sieved through a buffer soaked gel by an electrical current. Proteins are commonly separated in this manner using polyacrylamide gel electrophoresis (PAGE) to separate individual proteins in multiple samples or to examine multiple proteins within a single sample. There are several techniques available to study the variation of *Eimeria* proteomes including 1D and 2D SDS-PAGE (sodium dodecyl sulphate-polyacrylamide gel electrophoresis) (Vercauteren et al. 2007).

*Eimeria* species and strains have been studied to find any variation based on the proteome analysis. Barta et al. (1998) studied the intraspecific variation among five strains of *E. maxima* isolated in North America using one-dimensional SDS-PAGE. *Eimeria maxima* sporozoites showed at least 25 distinct bands for each strain but there were no detectable differences among the strains. In addition, 7 species of *Eimeria* infecting chickens have been studied using 2D SDS-PAGE to produce 'fingerprint' maps of the proteins from each species. All 7 species could be identified from their array of polypeptides but few differences were detected between strains of the same species (Sutton et al. 1989).
1.7.5 Amplified fragment length polymorphism (AFLP)

The principle of this technique is based on amplifying particular restriction fragments that are produced from genomic DNA. Subsequently, the amplified fragments are displayed by denaturing gel electrophoresis (Vos et al. 1995). The main steps of this technique are digestion of the genomic DNA with restriction enzymes and ligation of the resultant fragments with oligonucleotide adapters. The resultant fragments are then selectively amplified. Finally, the amplified fragments are separated using gel electrophoresis (Vos et al. 1995). Usually 50-100 bp fragments are amplified. A fingerprint for the analysed genome is created after it is separated by gel electrophoresis. Analysing genomic DNA using this technique does not require previous knowledge about the genome sequence, which means there is no need to design primers specifically targeting the genome of interest (Vos et al. 1995).

Polymorphic markers for *E. tenella* have been studied using AFLP to construct a genetic linkage map. Shirley and Harvey (2000) used different combinations of primers to study the same source of material. The best combination generated 10 polymorphic fragments per gel. The AFLP was also used to study nine samples of *E. maxima* (Houghton strain) using a series of primers and fingerprints that can be copied were generated (Blake et al. 2003).

This technique needs to employ combinations of restriction enzymes and primers that should be assessed to optimize the number of amplified fragments generated and the conditions of reaction. It also requires high quality and high molecular weight DNA. The disadvantage of using primers are not specifically designed to study a particular organism is that, the test can give false results that when contamination happens. Moreover, this technique cannot be employed to study mixed species infections because multiple sequences can be located and hidden within bands of the same molecular weight on the gel (Blake et al. 2003).
1.8 Transcriptome sequencing

The transcriptome is the complete set of RNA molecules (mRNA, rRNA, tRNA and non-coding RNA) transcribed by cells or organisms. The term is usually applied to a single cell or a population of similar cells at a particular time or stage of development under particular conditions. However, the term can also be used to refer to the whole set of transcripts that are predicted to be generated by a cell or an organism. Sufficient sequencing depth (the total number of sequencing reads) representing the cellular transcriptome can be achieved by using next generation transcriptome sequencing technologies such as 454 sequencing. In shotgun sequencing, cDNA is synthesized by reverse transcription of mRNA. Synthesized cDNA is fragmented and sequenced using pyrosequencing (e.g. 454 sequencing). Full length reads of a potentially large number of transcripts are available after the cDNA fragments are read, analyzed and contigs are generated (Morozova et al. 2009).

Pyrosequencing (454 pyrosequencing) has been used to sequence transcriptomes of plants and other biological systems. Libraries that have been generated and analysed using 454 sequencing include a number of plants such as Arabidopsis thaliana (Weber et al. 2007), Medicago truncatula and Zea mays (Cheung et al. 2006), as well as other biological systems such as Drosophila melanogaster (Torres et al. 2008), Caenorhabditis elegans (Shin et al. 2008) and human cell lines (Bainbridge et al. 2006). For example, the whole transcriptome shotgun sequencing (WTSS) procedure was used recently to survey the transcriptome of HeLa cells (Morin et al. 2008). The WTSS procedure is sometimes referred to as RNA-sequencing (Nagalakshmi et al. 2008).

Pyrosequencing is DNA sequencing based on the sequencing by synthesis. The principle of this technique is the detection of pyrophosphate (PPI), which is released as a result of nucleotide incorporation to synthesize DNA (Benkovic and Cameron 1995). Visible light is generated when one or more nucleotides are incorporated into a DNA molecule being synthesized from a cDNA template. The reaction is controlled by a cascade of enzyme reactions that lead to the release of light. The generated light can be detected by using a charge coupled device (CCD) camera, a photodiode or a photomultiplier tube. The
overall reaction time from polymerization to light detection takes place within 3-4 seconds (Ronaghi et al. 1996).

Pyrosequencing has been used to type complicated microbial communities using DNA markers that permit the study of both variable and conserved DNA regions. The variable region is usually sequenced using a DNA primer corresponding to the conserved or semi-conserved regions. Several bacterial species and strains have been differentiated by analyzing the 16S RNA gene. For instance, bacterial species and strains were taxonomically grouped based on investigating 20-100 nucleotides of the 16S rRNA gene using universal primers targeting the conserved sequences for amplification, followed by sequencing of the flanking variable region by one of the amplifying primers (Ronaghi 2001).

This method was used by Monstein et al. (2001) to identify and subtype variable V1 and V2 regions of Helicobacter pylori (14 S RNA). In another study, 16S rRNA was pyrosequenced by the same researchers to distinguish saprophytic bacteria and pathogenic bacteria from commensals found in the same host (Jonasson et al. 2002). However, longer read-lengths are essential to distinguish closely related bacterial strains by pyrosequencing. To address this issue, Gharizadeh et al. (2003) proposed the use of multiple group-specific sequencing primers to overcome the read-length restrictions.
Chapter 2 - Introduction and Hypotheses

2.1 Research Problem

The determination of immunologically important epitopes of coccidian parasites has been largely unsuccessful (Chapman et al. 2013) and, therefore, with one partially successful exception (e.g. CoxAbic) effective non-viable (subunit) vaccines against *Eimeria* species have not yet been produced despite the appearance of strong, protective immunity following a natural infection with these parasites.

2.2 Likely issues

Previous attempts to generate subunit vaccines against poultry *Eimeria* species have focused on exploiting potential immunogens originally selected using immune serum as a screen for ‘antigenicity’. Unfortunately, as discussed above, acquired immunity against *Eimeria* species infecting the domestic fowl is critically dependent on the generation of parasite-specific cytotoxic T-cells rather than plasma cells generating parasite-specific antibodies (Lillehoj 1986; Martin et al. 1993). The identification of immunologically relevant T-cell epitopes is technically more difficult than for B-cell epitopes. In addition, the complex eukaryotic genome of *Eimeria* species (Chapman and Shirley 2003) containing introns and exons, with potential for generating alternative transcripts through splicing events, prevents direct genome sequencing as a means of identifying species and strain-specific antigens. Finally, the large number of mRNA transcripts that have been sequenced from *Eimeria maxima*, *E. acervulina* and *E. tenella* (see http://www.coccidia.icb.usp.br/eimeriatdb/) have illustrated the existence of many Apicomplexa-specific transcripts that lack homology with known families of proteins (see Gajria et al. 2008 - http://toxodb.org/toxo/). For this reason, it is difficult to identify likely structural and/ or surface proteins that may be possible antigens of immunological importance.

The subject of this thesis is the comparison of two immunologically distinct strains of a single parasite species, *Eimeria maxima*. The Guelph and M6 strains of *E. maxima* have been shown to generate solid protective immunity in birds challenged with the homologous strain, yet birds immunized by natural infection with one strain remain
completely susceptible to infection with the second (Beattie et al. 2001). Other work showed that the sporozoite was targeted by the immune response upon secondary challenge infections. Despite being the target of protective immunity, there were no apparent differences in the sporozoite antigens recognized by immune chicken serum in each strain (e.g. Barta et al. 1997) and limited 2-dimensional SDS-PAGE of sporozoites of these two parasite strains showed few differences as well (J. R. Barta, pers. obs.); despite the virtually identical immune-serum antibody binding pattern in Western blots observed by Barta et al (1997), it is certainly possible that there were distinct, and potentially important, epitopes present that could not be detected in that study. Finally, even if the antibody binding was identical in both strains, it is unlikely that unique B-cell epitopes underlie strain-specific immune responses against *E. maxima* because bursectomized birds are capable of mounting a protective immune response in the absence of B-cells (i.e. no parasite-specific antibody production – see Rose and Hesketh 1979). Despite these observations, the strains illicit protective, strain-specific immune responses in chickens (via oocyst-induced infections) and therefore there ARE antigenic differences of importance between the strains.

### 2.3 Hypotheses

The foregoing observations have generated a number of hypotheses that will be tested in this thesis:

**Hypothesis 1:** A limited number of antigenic differences exist between the Guelph and M6 strains of *Eimeria maxima*;

**Hypothesis 2:** At least some of these antigenic differences are expressed on the sporozoite stage of the life cycles of these parasites;

**Hypothesis 3:** These limited antigenic differences in the sporozoites of the Guelph and M6 strains of *Eimeria maxima* are major determinants governing the lack of immunological cross reactivity observed between these parasite strains.
2.4 General Approach

To test these hypotheses, a direct comparison of the transcriptomes of these two closely related strains of the economically important chicken parasite *Eimeria maxima* was undertaken to identify (hopefully immunologically relevant) differences in the peptide sequences among putative immunogenic antigens expressed by these parasites as sporozoites. To obtain RNA samples from each strain that would be appropriate for analysis, peak oocyst production and kinetics of sporulation for each strain needed to be determined prior to the selection of suitable times for RNA collection from each strain. Once sampling times were determined, and appropriate parasites of each strain were collected, RNA isolation was accomplished.

The lack of a reference genome for *E. maxima* and the possibility that the protein(s) targeted by strain-specific immune responses could be relatively minor transcripts suggested that a comprehensive analysis of all sporulation/sporozoite transcripts be attempted. This meant that approaches such as expressed sequenced tags (EST) or screening of cDNA libraries generated from parasite mRNA that had previously been used to explore the transcriptomes of *Eimeria* spp. (Rangel et al. 2013; Shawarz et al. 2010) were unlikely to identify such subtle differences between strains. For these reasons, *de novo* sequencing of the entire sporogonic transcriptomes (early, mid and late sporulation as well as excysted sporozoites) of each strain was accomplished using a high throughput, next generation pyrosequencing platform.
Chapter 3 - The kinetics of oocyst shedding and sporulation in two immunologically distinct strains of *Eimeria maxima*, GS and M6

The contents of this chapter have been included in the following peer-reviewed publication:


3.1 Abstract

The kinetics of oocyst shedding and sporulation of two immunologically distinct strains of *Eimeria maxima* (GS and M6) were compared. Both strains had a prepatent period of approximately 120 h followed by peak oocyst shedding at 144-150 h post inoculation. Mean total oocyst output determined for each strain demonstrated that the fecundity of the M6 strain (12.8×10^3±1.95) of *E. maxima* was roughly twice that of the GS strain (6.9×10^3±3.33) when inoculated at the rate of 1,000 infective oocysts per bird. The process of oocyst sporulation was followed by repetitive sampling of sporulating oocysts at 26°C with aeration over a 138 h period. Sporulation was divided into five morphologically distinguishable stages whose abundance peaked at the following times during sporulation: unsporulated oocysts at 0 h; sporoblast anlagen at 18 h; sporoblasts without sporocyst walls at 22 h; and sporocysts without mature sporozoites at 38 h. The time to 50 % sporulation of *E. maxima* oocysts observed in the present study was approximately 53 h for both strains and all viable oocysts had completed sporulation by 60 h. In the present study, the prepatent periods, duration of oocyst shedding and the relative kinetics of sporulation of the GS and M6 strains of *E. maxima* were found to be virtually identical despite the immunological distinctiveness of these two parasite strains.

3.2 Introduction

Coccidiosis is a disease caused by obligate protozoan parasites belonging to the genus *Eimeria*. This disease infects the intestine of a host and reduces the growth efficiency. Coccidiosis is an important disease in poultry and other livestock (Lillehoj and Trout 1993; Shirley et al. 2007; Peek and Landman 2011). *Eimeria* species have complex life cycles that involve merogony and gametogony in the host, and sporogony in the environment. During sporogony, unsporulated oocysts develop into sporulated oocysts that have four sporocysts each containing two sporozoites (Hammond 1973). As many as seven morphologically distinct stages have been described during the sporulation of *Eimeria stiedai* (see Durr et al. 1971). For *Eimeria maxima*, six morphologically distinct stages were described by Tyzzer (1929) as follows: (a) unsporulated oocyst as discharged from intestine; (b) oocyst with band stretching through sporoplasm; (c) oocyst in stage just prior
to division (anlagen); (d) oocyst with four sporoblasts; (e) oocyst with immature sporocysts (sporocyst wall present but Stieda body not fully formed); and (f) fully mature oocyst with four mature sporocysts, each with a fully formed Stieda body and containing two fully formed sporozoites possessing refractile bodies. The Guelph strain of *E. maxima* (*E. maxima* GS) was initially isolated from a commercial poultry facility in Ontario, Canada, in 1973; a single-oocyst derived line of this isolate was then selected and has since been passaged at the University of Guelph. The M6 strain of *E. maxima* (*E. maxima* M6) was derived initially from a field isolate obtained from litter from a commercial broiler flock in Florida, USA in 1994 (see Martin et al. 1997). The M6 strain was selected by passage through birds that had been previously immunized against *E. maxima* GS. Single-oocyst derived lines were obtained, and one of these lines was designated *E. maxima* M6. The GS and M6 strains have been shown to be immunologically distinct and an infection with one strain does not provide cross-protective immunity against the other strain (Beattie et al. 2001). Both strains have been confirmed by a variety of molecular methods to belong to the same species, *E. maxima* (see Barta et al. 1998; Ogedengbe et al. 2011). As part of a comprehensive examination of the differences between these two immunologically distinct strains of *E. maxima*, GS and M6, the pattern of oocyst shedding for each strain was determined and then the kinetics of sporulation using oocysts obtained at the peak of oocyst shedding was characterized in detail.

### 3.3 Materials and methods

#### 3.3.1 Experimental animals and parasites

Day old Barred Rock chicks were obtained from Arkell Research Station (University of Guelph). They were raised in a coccidia-free facility until 28 days of age; experimental infections were initiated in 28-day-old chicks. Birds were provided a 12 h/12 h light–dark cycle and provided feed and water ad lib; experimental procedures were carried out in compliance with the Canadian Council on Animal Care guidelines and have been approved by the University of Guelph's Animal Care Committee. *E. maxima* oocysts (strains GS and M6) were propagated separately in chickens raised coccidia-free in the University of Guelph's Central Animal Facility Animal Isolation Unit. Oocysts were collected by salt flotation using standard methods (Ryley et al. 1976). The partially purified oocysts were
suspended in 2.5 % potassium dichromate (w/v, aqueous) before incubation on a rotary
platform shaker at 26 °C to sporulate for a minimum of 72 h. All parasites for experimental
infections were stored at 4 °C for less than 60 days prior to use. Immediately prior to
inoculation into experimental birds, oocysts were collected by centrifugation (1,000×g for
10 minutes) and then suspended in distilled water; sporulated oocysts were counted and
diluted to the appropriate dosage with additional distilled water.

3.3.2 Determining the peak of oocyst shedding

Ten 28-day-old chickens were divided randomly into two groups. Each chicken was
placed in a separate wire-floored cage. Chickens were inoculated by oral gavage with
1.0×10³ oocysts per bird using a 1-cc tuberculin syringe (no attachment fitted). Five of the
chickens were inoculated with the M6 strain and the other five chickens were inoculated
with GS strain. Fecal samples were collected continuously throughout days 6–10 post
inoculation (PI). A stack of 16 foil sheets that covered the entire cage floor were placed
beneath infected birds at the end of day 5 post inoculation. All feces from each bird were
collected from 120 to 216 h post inoculation (HPI) in 6 h increments by removing the
topmost sheet of foil each 6 h; the foil held all feces that had accumulated over the most
recent 6 h period. Contaminants such as food and feathers were removed from the collected
fecal samples that were then placed into individual sterile beakers. The total number of
oocysts shed during each 6 h interval was determined using a McMaster counting chamber
(Appendix 3.1). Oocysts were allowed to float clear of debris for 5 minutes before
counting using a light microscope fitted with a long focal length 16× objective. The
following equation was used to calculate the total number of oocysts in a fecal sample:

Total oocysts per bird

\[
= \frac{(\text{total sample volume (ml)} \times \text{oocyst count} \times \text{dilution factor})}{(0.15 \text{ ml (grid volume)} \times \text{number of grids counted})}
\]

Total oocyst output per bird was calculated by summing the counts from all samples
from a single bird. Parasite fecundity in each bird was then calculated as the total oocyst
output divided by the inoculating dose. Mean total oocyst output and mean parasite
fecundity of the GS and M6 strains of *E. maxima* were compared using Student's t test assuming unequal variances and were considered statistically different if p<0.05.

### 3.3.3 Kinetics of oocyst sporulation

Ten 28-day-old chicks were divided into two groups and half of the chickens were infected orally with 2.0×10⁴ oocysts of *E. maxima* GS and the remainder were infected orally with *E. maxima* M6. Fecal samples were collected for a 6 h period spanning the time of peak oocyst shedding determined previously for each strain. Feces containing oocysts from each strain (GS and M6) were blended with potassium dichromate (2.5 %w/v aqueous) in separate 500-ml Erlenmeyer flasks and then incubated on the same rotary shaker at 26 °C to sporulate. A sample was taken from each flask every 4 h (from 0 to 138 h) to document the morphological changes during sporulation. Five morphologically distinct stages of sporulation could be identified reliably in sporulating oocysts of *E. maxima* (Figure 3.1). Thus, immediately after collection at each time point, 100 viable oocysts from each fecal sample were examined using an Olympus Provis AX70 microscope (Olympus Canada, Richmond Hill, ON). Each oocyst was assigned to one of five developmental categories based on its morphological appearance: (1) unsporulated; (2) sporoblast anlagen; (3) sporoblasts without sporocyst walls; (4) sporocysts without mature sporozoites; and (5) sporocysts with mature sporozoites (Appendix 3.2 and Appendix 3.3)
Figure 3.1 Oocysts passed through 5 morphologically identifiable stages during sporulation: a) unsporulated oocyst with single spherical zygote within an oocyst wall (0 h sporulation); b) sporoblast anlagen formation (18h sporulation); c) and d) separate sporoblast but no evidence of sporocyst walls (20-24 h sporulation); e) sporocyst walls formed but sporozoites not yet fully formed (38 h sporulation); f) fully formed oocyst with mature sporozoites within sporocysts with fully formed Stieda bodies (> 60 h sporulation). The kinetics of sporulation through these stages is summarized in (Figure 3.4).

3.4 Results

3.4.1 Comparison of oocyst shedding of E. maxima GS and M6 strains

Both strains began shedding unsporulated oocysts during the 120-126 HPI sampling period that corresponds to a prepatent period of about 5 days. For the majority of infected birds, peak oocyst shedding was found to occur during the 144-150 HPI sampling period for both the M6 strain (for three of five birds, Figure 3.2) and the GS strain (for four of five birds, Figure 3.2). The mean oocyst output for each strain at each 6 h collection period is illustrated in (Figure 4.3); there was no difference between the two strains with respect to the time PI at which peak oocyst output occurred. Oocysts continued to be shed for the duration of the sampling period, albeit at low numbers from 198 HPI onward. Birds inoculated with 1,000 oocysts of E. maxima M6 shed more oocysts than birds infected with the GS strain in all sampling periods except one. Birds infected with the M6 strain produced significantly more oocysts than birds infected with 1,000 oocysts of E. maxima.
GS (p<0.05) during the 144-150 HPI collection period. This sampling period corresponded to the peak of oocyst shedding (5.96×10^6 versus 2.82×10^6 oocysts/bird, respectively). Significantly, more (p<0.05) oocysts were shed by the E. maxima M6-infected birds during the sampling periods of 192-198, 198-204, and 204-210 HPI. This increased productivity was also reflected in the significantly higher mean total number of oocysts shed by birds infected by E. maxima M6 (p < 0.05) over the course of a single infection (12.76×10^6 oocysts/bird, n05) compared with the mean total number of oocysts shed by birds infected by E. maxima GS (6.86×10^6 oocysts/ bird, n05). The mean fecundity (number of progeny oocysts produced per oocyst inoculated) of the E. maxima M6 (12.8×10^3±1.95) was significantly higher (p<0.05) than the mean fecundity for E. maxima GS (6.9×10^3±3.33).

3.4.2 Determination of the kinetics of oocyst sporulation

The sporulation process was divided into five morphologically distinct stages that were then enumerated in a mixture of sporulating oocysts: (1) unsporulated (Figure 3.1a); (2) sporoblast anlagen (Figure 3.1b); (3) sporoblasts without sporocyst walls (Figure 3.1c, d); (4) sporocysts with sporocyst walls but lacking mature sporozoites (Figure 3.1e); and (5) sporocysts with mature sporozoites (Figure 3.1f). The kinetics of sporulation of E. maxima M6 at 26 °C is illustrated in Figure 3.4. The first group is unsporulated oocysts (at 0 h) that developed to sporoblast anlagen stage at 22 h. The sporoblast anlagen then developed to sporoblast without sporocyst wall at 24 h. As the incubation of oocysts continues, the sporoblast without sporocyst developed to next stage which is sporocysts without mature sporozoites at 38–40 h. The final stage is the sporocysts with mature sporozoites at 60 h and above. The kinetics of GS strain is shown in Figure 3.4. It is evident from the two figures that there is no difference in the pace of sporulation between the two strains of E. maxima (M6 and GS) through all of the five enumerated stages (unsporulated, sporoblast anlagen, sporoblasts without sporocyst walls, sporocysts without mature sporozoites, and sporocysts with mature sporozoites).
Figure 3.2 Pattern of oocyst shedding by *Eimeria maxima* GS and M6 in chicks inoculated with 1000 oocysts/bird at time 0. Total oocyst output during each 6 h collection period for 5 birds (one line plotted per bird) for *E. maxima* GS (top) and *E. maxima* M6 (bottom). The peak of oocyst output was during the 140-150 HPI collection period for both strains.
Figure 3.3 Mean total oocyst output of *Eimeria maxima* GS and *E. maxima* M6 in chicks inoculated with 1000 oocysts/bird at time 0 sampled continuously in 6 h collection periods from 120 to 216 HPI. Columns represent the mean total oocyst output (SD indicated above each bar) from 5 individual birds inoculated with the indicated parasite strain during each 6 h collection period. Mean total oocyst counts during a single collection period that differ statistically (p<0.05) between *E. maxima* GS and *E. maxima* M6 are indicated by an asterisk.
Figure 3.4 Kinetics of oocyst sporulation at 26°C of *Eimeria maxima* GS (top) and *E. maxima* M6 (bottom) were followed by determining the proportion of 100 viable oocysts at each of five stages of parasite development during sporulation (Figure 3.1). At the start of sporulation only unsporulated oocysts were present (dotted lines) but these quickly developed to the sporoblast anlagen form that peaked in abundance at 22 h (dashed lines). Sporoblasts without sporocyst walls appeared shortly thereafter and peaked in abundance at 24 h (hollow dashed lines). Sporocysts possessing walls but without mature sporozoites first appeared at 28 h and peaked in abundance at 44 h (hollow lines). Fully mature oocysts containing sporocysts with mature sporozoites first appeared at 48 h and sporulation was complete by 60 h (solid lines). There were no differences observed between the sporulation kinetics of *E. maxima* GS and *E. maxima* M6.
3.5 Discussion

The process of sporulation of eimeriid coccidia is a complex and necessary feature of all coccidian life cycles; nonetheless, the process of sporulation has been explored infrequently. As part of a comparison of two immunologically distinct strains of *E. maxima* (GS and M6), the timing of oocyst shedding and kinetics of sporulation were compared between the two strains under the same environmental conditions. The prepatent period of *Eimeria* species has been used as one of the distinguishing features to help identify species in a host. The peak of oocyst shedding was at 144–150 HPI (between days 5 and 6 postinoculation) in both strains of *E. maxima* (GS and M6). In the original description of *E. maxima* by Tyzzer (1929), chicks receiving large numbers of infective oocysts began to shed small numbers of unsporulated oocysts at the end of 6 days PI (~156 h) and continued shedding oocysts until about 10 days PI; it was clear from Tyzzer's description that in birds given lower doses of oocysts, his strain of *E. maxima* had a prepatent period of about 7 days (168 h). In contrast, both the GS and M6 strains of *E. maxima* demonstrated a shorter prepatent period (120–126 h) than Tyzzer's *E. maxima* strain. The observations on the prepatent period of these *E. maxima* strains in the present study agree with the 121 h minimum prepatent period listed by Reid and Long (1979) and 123 h determined by Edgar (1955) with their strains of *E. maxima*. Despite the utility of the prepatent period helping to identify *Eimeria* species of domestic fowl obtained from recent field isolates, the prepatent period can be significantly altered by selective passage in a process known as attenuation. Parent lines of *E. maxima* could have their typical prepatent period of about 120 h reduced to less than 107 h (McDonald et al. 1986). According to these authors, this abbreviated prepatent period is the result of a reduced number of merogonic cycles occurring before gametogony and oocyst formation. Both the M6 and GS strains of *E. maxima* have been passaged in a laboratory environment for some time with oocyst collection generally performed during day 6 PI; it is possible that such repetitive passaging has selected for a slightly shorter prepatent period than in strains cycling naturally among birds.

Although the timing of oocyst shedding did not differ between the strains, the fecundity of the strains differed significantly. Fecundity, or replicative potential, represents the number of oocysts produced per oocyst inoculated into a suitable and immunologically
naïve host (Johnston et al. 2001). Fecundity is calculated by dividing the total number of oocysts shed by a bird by the number of oocysts in the inoculating dose. The number of oocysts produced from chickens that were infected with M6 strain (mean fecundity of $12.8 \times 10^3 \pm 1.95$) was almost exactly double the number of oocysts that were produced from chickens infected with GS strain (mean fecundity of $6.9 \times 10^3 \pm 3.33$) with the same dose of oocysts ($1.0 \times 10^3$ oocysts/bird). Fecundity of various *Eimeria* species is related inversely to the infectious dose of oocysts provided to birds; larger inoculating doses may produce more oocysts per bird during the course of an infection but the fecundity per oocyst inoculated drops off rapidly (Johnston et al. 2001; Williams 2001). Assuming that the M6 and GS strains of *E. maxima* had the same infectivity, the observed difference in fecundity between the strains can be explained by a single additional mitotic division during the merogonic portion of the life cycle of the M6 strain compared with the life cycle of the GS strain. Allen et al. (2005) suggested that increased fecundity of an *E. maxima* strain was associated with increased cross-immunity to other *E. maxima* strains. Although not examined directly in this study, *E. maxima* M6 was derived from a strain of coccidia that was somewhat more cross protective against the GS strain than vice versa (Beattie et al. 2001) which is in agreement with this suggestion by Allen et al. (2005). However, transcriptome analyses of sporulating oocysts of the *E. maxima* GS and *E. maxima* M6 strains have uncovered numerous differences in expressed GPI-anchored and other surface antigens (see Chapter 5). These differences could explain the previously demonstrated strain specific immune responses against these *E. maxima* strains (Beattie et al. 2001). Whether or not increasing fecundity, on its own, is sufficient to elicit improved cross-protective immune responses to other strains of *E. maxima* remains to be determined. Sporulation of oocysts occurs outside the host in the environment and takes about 2–3 days as the unsporulated oocysts develop into infective, sporulated oocysts. Unsporulated oocysts shed into the environment with feces contain a diploid zygote. Ultrastructural studies demonstrate that diploid oocysts undergo classical meiosis with the two divisions producing four haploid sporoblasts (Ferguson et al. 1978). Following meiosis and sporoblast formation, a single mitotic division within each sporoblast gives rise ultimately to four mature sporocysts each containing two sporozoites (Canning and Anwar 1968; Fergusson et al. 1978; Beesley and Latter 1982; Kinnaird et al. 2004). The general
sporulation process was described by Tyzzer (1929) in which he identified morphologically distinct stages of sporulation of oocysts of *Eimeria* spp. from gallinaceous birds. These were: (a) oocyst as discharged from intestine; (b) oocyst with band stretching through protoplasmic mass; (c) oocyst in stage just prior to division; (d) oocyst showing division into four cells, or sporoblasts; (e) oocyst, stage showing immature spores with membrane slightly developed; and (f) oocyst, spores fully developed, each with its two sporozoites fully formed. These developmental forms were similar to our recorded developmental stages except that we combined Tyzzer's (1929) stages “a” and “b” into a single “unsporulated” form characterized by a single, spherical sporoplasm within the oocyst. In all other regards, our observations and illustrations agree with Tyzzer's (1929) original description and illustrations of the sporulation of *E. maxima*.

Norton and Chard (1983) determined the oocyst sporulation times of seven *Eimeria* species from the domestic fowl. The time at which 50 % of the oocysts were sporulated for each *Eimeria* species was documented as follows: 11.4 h for *Eimeria acervulina*; 19.0 h for *Eimeria mivati*; 19.7 h for *Eimeria necatrix*; 21.2 h for *Eimeria tenella*; 24.8 h for *Eimeria praecox*; 38.1 h for *Eimeria maxima*; and 38.3 h for *Eimeria brunetti*. The time for 50% of *E. maxima* oocysts to sporulate in the present study (approximately 53 h) was considerably longer than the 38 h that Norton and Chard (1983) documented for their strain of *E. maxima*. However, sporulation time is sensitive to the temperature at which the oocysts are held during sporulation (Edgar 1954). Sporulation was completed at 26 °C in the present study whereas Norton and Chard (1983) sporulated their oocysts at 29 °C. In addition, sporulation is an aerobic process and consequently sporulation time is affected by available oxygen in the sporulation medium (Marquardt et al. 1960); this is usually addressed by aeration, by forcing air through the sporulation medium (Canning and Anwar 1968), or through agitation such as the platform rotator used in the present study. In earlier studies, sporulation was achieved with oocysts stored in dichromate solutions within Petri dishes without agitation or aeration; this makes direct comparisons between trials difficult. In the environment of the poultry house, oocysts must obtain sufficient oxygen for the sporulation process as well. Waldenstedt et al (2001) determined that sporulation of *E. maxima* was reduced in the samples with the highest moisture content (62 %) and increased under the driest litter conditions studied (16 % moisture content). The authors concluded that drier
litter conditions favoured sporulation because of the increased availability of oxygen in the driest substrates compared to the limited amount of oxygen available in moister substrates. Finally, the subjective criteria used to designate oocysts as “fully sporulated” morphologically were not necessarily the same in both studies. In the present study, the fecundity of *E. maxima* M6 was determined to be approximately twice that of *E. maxima* GS. However, the prepatent periods, duration of oocyst shedding, and the relative kinetics of sporulation of *E. maxima* M6 and GS were found to be virtually identical despite the immunological distinctiveness of these two parasite strains.
Chapter 4 - Functional annotation and mapping of transcript sequences of two immunologically distinct strains (Guelph and M6) of the chicken coccidium *Eimeria maxima*

4.1 Abstract

Two immunologically distinct strains of poultry coccidia, *E. maxima* GS and *E. maxima* M6, were investigated in this study. First, total RNA samples were collected from four stages of sporozoite development (18 h, 22 h, 38 h and fully sporulated). Messenger RNA was obtained in order to synthesize cDNA for each strain. The cDNA samples were then sequenced using Illumina. The resulting sequencing data (size of read is 100 bp) were paired and *de novo* assembled using Geneious. Ten thousand transcripts from both strains were blasted against NCBI using Blast2GO. The results of this study showed that 2067 transcripts of GS strain and 1610 transcripts of M6 strain were assigned putatively to biological function(s). The majority (~60%) of transcripts for which putative function could be assigned in both strains were assigned to metabolic and cellular process. Despite their lack of immunological cross-reactivity, the GS and M6 strains of *E. maxima* expressed quite similar transcriptomes; of the 5,000 transcripts per strain that were examined, only 4.23% of transcripts expressed by one strain (193 transcripts from GS and 230 transcripts from M6) did not appear to have a homolog expressed by the other strain. These 423 transcripts may be fertile starting points for identifying strain-specific molecules. A number of transcripts encoding putative surface antigens (SAgs) were identified among the generated transcripts: 19 SAg transcripts were identified in the GS strain and 18 different SAg identified in the M6 strain. All identified SAgs were further analyzed in chapter 5.

4.2 Introduction

Coccidiosis is an economically important intestinal disease of poultry and other animal species. This disease is caused by seven *Eimeria* species in chickens causing losses in poultry industry that exceeds £2 billion annually (Shirley et al., 2005). The conventional methods that have been available to control coccidiosis are mainly using prophylactic chemotherapy (anti-coccidial drugs) and vaccination with live vaccines. These two methods have been effective in controlling coccidiosis, but, drawbacks include drug resistance and the limitations with producing live vaccine are not unimportant. From above, the search for a new and safe approach to control coccidiosis is needed.

Two immunologically distinct strains, of *Eimeria maxima*, GS and M6 (see Beattie et al., 2001) were investigated in this study. Despite their relative lack of immunological cross-reactivity, these two strains are similar in their kinetics of oocyst sporulation and
oocyst production (Al-Badri and Barta 2012). The Guelph strain of *Eimeria maxima* (*E. maxima* GS) was isolated from an Ontario, Canada, poultry facility in 1973 and a single-oocyst derived line has been maintained at the University of Guelph since then. *Eimeria maxima* M6 was derived from parasites initially obtained on a Florida, USA, broiler farm in 1994. Each of these two strains has been characterized extensively (Martin et al 1997; Barta et al. 1998; Beattie et al. 2001; Al-Badri and Barta 2012).

The first asexual stage of the *Eimeria* life cycle, the sporozoite is known to be a target of the immune system (e.g. Beattie et al. 2001) but there has been ambiguity about the sporozoite antigenic molecules that stimulate the immune system of chickens. Jeffers and Long (1985) found that sporozoites of *E. tenella* inhibited with the anticoccidial drug decoquinate stimulated significant protective immunity in chickens subsequently challenged with decoquinate resistant *E. tenella* sporozoites. The elicited immunity eliminated the decrease in weight gain but provided only slight protection against intestinal lesions. Also, Jenkins et al. (1991) found that protective immune responses were elicited only after invasion by *E. tenella* sporozoites. Although considerable sequence information is known from *E. tenella* and from the merozoites of *E. maxima* and other species, no data are available regarding the expression of genes by *E. maxima* during this critical period of development of the parasite.

The objective of this study was to determine the potential biological function(s) of transcriptomic sequences generated from *E. maxima* GS and *E. maxima* M6. The sequencing data obtained using Illumina could provide new insights into the transcriptomes of these two immunologically distinct *E. maxima* strains during the critical phase of the life cycle of these coccidia during which sporozoites are formed, mature and get activated prior to the penetration of host enterocytes. These excysted sporozoites and the structural molecules that comprise them are demonstrated targets of a protective immune response against *E. maxima* (see Beattie et al. 2001) and so any differences observed may help to explain the lack of immunological cross-protection between these parasite strains. In particular, these data may also be useful for studying the differences in surface antigens (SAgs) between these two immunologically distinct strains (GS and M6).
4.3 Materials and methods

4.3.1 Parasite Propagation

Twenty day-of-age Barred Rock chicks were obtained from Arkell Research Station (University of Guelph). They were reared coccidia free in the Animal Isolation Unit of the Central Animal Facility, University of Guelph to 28 days-of-age. Experimental infections were initiated in 28 day-old chicks. Animals were provided a 12 h/12 h light-dark cycle and provided feed and water ad lib. All procedures were carried out in compliance with the Canadian Council on Animal Care guidelines and have been approved by the University of Guelph’s Animal Care Committee.

At 28 days-of-age, birds were transferred to disposable caging containing a suspended wire mesh floor. Half of the chicks were inoculated with $2 \times 10^4$ oocysts/bird of *Eimeria maxima* M6 and the remainder were inoculated with the same dose of *Eimeria maxima* GS via oral gavage using a sterile 1cc tuberculin syringe without a needle fitted. Aluminum foils were set in bottom of the boxes at 5 days post inoculation. The foils remained for 6 h to collect feces. Based on the previous result (Al-Badri and Barta 2012), the peak of oocysts production is at 150 h post-inoculation (HPI). The foils were set on 144 HPI and were left for 6 h to collect the feces. At 150 HPI, fecal samples were processed according to Al-Badri and Barta (2012). Oocysts were then concentrated using a salt flotation technique (Ryley 1976; Al-Badri and Barta 2012).

4.3.2 Oocysts sporulation and bleaching

Unsporulated oocysts partially purified by salt flotation were then surface sterilized using ice-cold household bleach (4.25% sodium hypochlorite) prior to sporulation so that subsequent samples obtained during the sporulation process would be free of exogenous RNA from fecal microorganisms and could be frozen for RNA extraction without further processing. Sporulation of surface sterilized oocysts was accomplished in 2% potassium dichromate (w/v aqueous) at 26°C on a rotary platform as previously described (Al-Badri and Barta 2012, see Chapter 3).
4.3.3 RNA preparation

4.3.3.1 Samples used for RNA extraction

Previous work had identified three time points during sporulation (18 h, 22 h, and 38 h) that corresponded to sporoblast analagen, sporoblasts without sporocyst, and sporocysts possessing walls but without mature sporozoites, respectively (see Al-Badri and Barta 2012), and fully sporulated sporozoites undergoing excystation. The middle of the increasing 'stage' frequencies was used to catch the mRNA that was being generated to build the stage of development under construction. For example, although partially sporulated oocysts that had sporocyst walls but no mature sporozoites were found at 44 h maximally, a time just before then was used to collect mRNA. At each time point, partially sporulated oocysts were washed with ice cold water 3 times to remove excess potassium dichromate, pelleted and then stored at 80°C.

For obtaining excysting sporozoite RNA, sporocysts from fully sporulated oocysts were isolated by breaking the oocyst walls using a Mickle disintegrator (Mickle Laboratory Engineering Co. Ltd. Gomshall, UK) containing 0.5 mm silica beads. Sporocysts were separated from oocyst walls and unbroken oocysts by sieving through Nitex screen-printing cloth with a 10 μm pore size (Fernando and Pasternak 1977). The filtrate containing sporocysts was excysted by incubation in 0.25% trypsin (Sigma, St. Louis, MO) and 5% (v/v) chicken bile at 41°C for ~2 h using a shaking water bath. At the end of the incubation, excysted sporozoites were collected by centrifugation and then stored immediately at -80°C.

4.3.3.2 Extraction of RNA

Partially sporulated oocysts (18 h, 22 h, and 38 h) needed to be bead disrupted prior to RNA isolation. Briefly, thick slurry of concentrated oocysts in a minimal amount of 1× PBS (~400 µl) was transferred to a 1.5 ml RNase-free microcentrifuge tube to which the same volume of autoclaved glass beads (~2.3 g) was added. The sample was then shaken for 5 minutes at high speed using a Mickle disintegrator (H. Mickel Laboratory Engineering Co., Gomshall, UK) to completely disrupt the oocysts, sporocysts and sporozoites to release RNA.
Ribonucleic acid extraction from bead disrupted, sporulated oocyst samples (18 h, 22 h and 38 h) and sporozoites was accomplished using the FastRNA® Pro Green Kit (MP Biomedicals, Santa Ana, CA, USA) according to the manufacturer’s instructions with the exception that no further homogenization steps were performed beyond the Mickle disintegrator treatment above. At the conclusion of the protocol, pellets of RNA were resuspended with 20 µl of DEPC-treated water. The concentration and relative purity (OD\textsubscript{230}/OD\textsubscript{260} ratio) of the dissolved RNA was measured spectrophotometrically using a Nanodrop 2000 (Thermo Scientific, DE, USA). Ribonucleic acid samples were further cleaned and desalted using RNeasy® MinElute® kit (QIAGEN, Valencia, CA). Four RNA samples (18 h, 22 h, 38 h, and sporozoites) were adjusted to a volume of 100 µl with RNase-free water and processed according to the kit manufacturer’s manual. The resulting RNA samples were then quantitated and stored at -80°C prior to their use in preparing a cDNA library from each strain.

4.3.3.3 Further RNA purification

Ribonucleic acid samples were further cleaned and desalted using RNeasy® MinElute® kit (QIAGEN, Valencia, CA). In the beginning of this trial, RNA samples were adjusted to a volume of 100 µl with RNase-free water. Buffer RLT (350 µl) was added to the diluted total RNA samples and was then mixed well. Ethanol (96-100%, 250 µl) was then added to the dilute RNA samples. Each sample (700 µl) was then transferred to an RNeasy MinElute spin column placed in a 2 ml collection tube. The lid was closed tightly and the tube was centrifuged for 15 seconds at 10,000 rpm; and the flow through was discarded.

The RNeasy MinElute spin column was placed in 2 ml collection tube and Buffer RPE (500 µl) was added to the spin column immediately. The lid of the spin column then was closed gently and tightly. The tube was centrifuged for 15 seconds at 10,000 rpm to wash the spin column membrane. The high speed centrifugation forced the solution to pass through and the flow through was discarded. Eighty percent ethanol (500 µl) was added to the RNeasy spin column. The lid was closed gently, and the tube was centrifuged for 2 minutes at 10,000 rpm to wash the spin column membrane.
The flow through and collection tube were discarded. Then, the RNeasy MinElute spin column was placed in a new 2 ml collection tube. The lid of the spin column was left open and the tube centrifuged at full speed for 5 minutes. The flow through and collection tube were discarded. Finally, the RNeasy MinElute spin column was put in a new 1.5 ml collection tube and RNase-free water was directly added to the center of the spin column membrane. The lid was closed gently, and the tube centrifuged for 1 minute at full speed to elute the total RNA.

4.3.3.4 RNA quality assessment using Formaldehyde Gel Electrophoresis

Total RNA samples were assessed using formaldehyde submarine agarose gel electrophoresis to check for integrity of large and small subunit rRNA. Typically, a small size gel (10 × 6.5 cm) with 1% agarose was used. Agarose (0.5 g) was measured and mixed with 50 ml 1× MOPS running buffer (Appendix 4.1). The mixture was put into a 100 ml Erlenmeyer flask and heated to near boiling until the agarose dissolved. After the agarose solution cooled to ~55 °C, 0.9 ml of formaldehyde (37% w/v) was added and then the formaldehyde-agarose solution was immediately poured into a prepared gel casting tray in a fume hood and was left there to harden for 30 minutes. The gel was then placed into a gel box and submerged in 1× MOPS running buffer.

Ribonucleic acid samples were denatured by heat to disrupt the secondary structure of RNA. Each total RNA sample (1-3 µg) was mixed with 0.5× volume of formaldehyde load dye. Samples were heated at 70°C for 5-10 minutes to denature. Samples were then immediately cooled with ice to prevent the secondary structure from reforming. Samples were centrifuged briefly and loaded immediately onto the gel. Deoxyribonucleic acid ladder (100 bp) was used as a reference. Finally, the gel was run at 5 – 7 V/cm until bromophenol blue dye front was 2/3 of the way through the gel.

4.3.4 Amplification of TRAP gene (PCR)

Complementary DNA quality was tested by amplifying the thrombospodin-related adhesive protein (TRAP) gene that is known to have a long (~7990 bp) mRNA (c.f. GenBank AY239227). For the purposes of this test, cDNA was synthesized from total RNA extracted from 22 h partially sporulated oocysts using a SuperScript Double-Stranded
cDNA Synthesis Kit (Invitrogen, Grand Island NY) according to the manufacturer’s instructions. Briefly, synthesis of the first-strand cDNA from total RNA samples was initiated by mixing oligo dT primer (100 pmol), 30 µg total RNA, and 10 µl DEPC-treated water. This mixture was heated to 70°C for 10 minutes; and quickly chilled on ice. Contents of the tube were collected by brief centrifugation.

Subsequently, 4 µl 5x first-strand buffer, 2 µl MDTT (0.1), and 1 µl (10mM dNTP mix) were added. The mixture was gently vortexed and the reaction collected by brief centrifugation. The tube was placed at 45°C for 2 minutes to equilibrate the temperature. Superscript II RT (1 µl) was added and the reaction was mixed gently. The tube was then incubated at 45°C for 1 h. The tube was placed on ice to terminate the reaction.

The second-strand of cDNA was synthesized from the same reaction. On ice, the following reagents (91 µl DEPC-treated water, 30 µl 5x second strand reaction buffer, 3 µl of 10mM dNTP mix, 1 µl E. coli DNA ligase (10 U/µl), 4 µl E. coli DNA polymerase I (10U/µl), 1 µl E. coli RNase h (2 U/µl)) were added to the reaction. The mixture was gently shaken (vortexed) and then incubated for 2 h at 16 ºC. Two µl of T4 DNA polymerase were added, and samples then were incubated at 10°C for 5 minutes.

The tube was then placed on ice and 10 µl of 0.5 M EDTA were added to the completed second strand reaction. Phenol:chloroform:isoamyl alcohol 25:24:1 (160 µl) was added and samples then thoroughly vortexed and centrifuged at room temperature for 5 minutes at 14,000×g to separate the phases. The upper, aqueous layer (140 µl) was carefully removed and transferred to a fresh 1.5 microcentrifuge tube. Ammonium acetate (7.5M NH₄OAc) 70 µl was added, followed by 0.5 ml of ice-cold absolute ethanol. The mixture was then vortexed thoroughly and immediately centrifuged at room temperature for 20 minutes at 14,000×g.

After centrifugation, the tubes were removed and the supernatant was carefully discarded. The pellet was overlaid with 0.5 ml of ice-cold 70% ethanol. The tubes were then centrifuged for 2 minutes at 14,000×g; the supernatant was discarded. Complementary DNA was dried at 37°C for 10 minutes to evaporate residual ethanol; and finally the pellet
was dissolved in a small volume of DEPC-treated water (3 µl per 25 µg of starting total RNA); and stored at - 80°C.

The PCR reaction used ~100 ng ds cDNA, 10 µl Master Mix, and 2.5 µl of primers (TRAP-EP006 5’-CCGAATTGCACCCCA-3’ and TRAP-EP007 5’-CTGAATGTCCCG CTGTGC-3’, see Witcombe et al. 2003). These primers were designed to amplify a 647bp product (nt 6664 to 7310 of AY239227, i.e. starting about 1,300 bp from 3’ end of the mRNA). The reaction was then run for 40 cycles in a thermal cycler programmed with the following temperature profile: denaturation at 94 °C for 5 minutes, annealing 60 °C for 15 seconds, and extension at 72 °C for 60 seconds.

The resulting PCR product was electrophoresed using a 1% agarose submarine gel that was prepared by mixing 0.5 g agarose (Mandel Scientific, Guelph, Ontario, Canada) and 50 ml of 1× TBE. The mixture was put into a 100 ml Erlenmeyer flask and heated to near boiling until the agarose dissolved. After the agarose solution cooled to ~55 °C, 1 µl of ethidium bromide (10mg/ml, w/v aqueous) was added. The mixture was immediately poured into a prepared gel casting tray and was left to harden for 30 minutes. The gel was then placed into an electrophoresis unit and submerged in 1×TBE running buffer. The DNA ladder was 100 bp. The gel was loaded with sample and DNA size ladder and run for approximately 1 h until the tracking dye was nearing the edge of the gel. Bands were visualized using UV transillumination and the size of any products estimated by comparison with DNA size standards.

4.3.5 Complementary DNA Library Construction and Next Generation Sequencing (Illumina)

Equal amounts of RNA (1µg) were taken from each bulk cellular RNA sample (18 h, 22 h, 38 h and excysting sporozoites) to make a single pooled RNA sample (4 µg) for each of the M6 and GS strains of Eimeria maxima. Messenger RNA was isolated from bulk cellular RNA from each pooled sample (from a single strain) using the TruSeq RNA Sample Preparation Kit (Illumina Inc, San Diego, CA) according to the manufacturer’s instructions. The resulting mRNA was then fragmented to generate pieces with a mean length of approximately 250 bp. The fragmented mRNA was then converted to cDNA,
adenylated on the 3’ end before linking to adaptors unique to each *E. maxima* strain. *Eimeria maxima* GS was linked to adaptor 1 (CTTGTA) and *E. maxima* M6 was linked to adaptor 2 (GCCAAT). After adaptor linkage, the resulting cDNA libraries were pooled prior to paired-read sequencing at 100bp per read using an Illumina HiSeq 2000 Next-Generation Sequencer at The Centre for Applied Genomics (TCAG), University of Toronto (Toronto ON, Canada).

### 4.3.6 De novo assembly and Blast2GO analysis

Initial *de novo* assembly of 100 bp paired reads generated by Illumina sequencing was accomplished from 8×10⁶ reads (approximately 800 Mb of sequence data/strain) from each of *E. maxima* M6 and *E. maxima* GS using the *de novo* assembly function within Geneious Pro (version 6 and later, Drummond et al. 2012). The *de novo* assembly parameters were as follow: Maximum ambiguity (4); Index word length (10); Maximum gap size (1); Maximum matches per read (10%); paired read distance was not restricted to nearby mappings to improve assembly likelihood. At the conclusion of each *de novo* assembly, 10,000 transcripts were retained. For each *E. maxima* strain, consensus sequences were calculated from each of the 10,000 assembly transcripts based on read quality and these 10,000 consensus sequences formed the transcript dataset for annotation and further analyses.

Preliminary identification and annotation of the most abundant 5,000 transcripts (based on number of reads associated with each transcript) from each strain was accomplished using Blast2GO’s Blast, Mapping and Annotation functions ([www.blast2go.com](http://www.blast2go.com)), (see Conesa et al. 2005). Briefly, for each *E. maxima* strain, the 5,000 most abundant transcript output from the *de novo* assembly were searched against the non-redundant (nr) sequence database (accomplished using Blast2GO version 2.4.6 and the following search parameters: Blastx search algorithm; non-redundant protein sequence database (nr, comprised of all non-redundant GenBank CDS translations + PDB + SwissProt + PIR+ PRF excluding environmental samples from WGS projects); and discriminating hits using an expect (E) value of ≤1.0E-3. After the 5,000 Blast searches of individual transcripts were completed, mapping was carried out to link Blast hits to gene oncology (GO) databases including the GO ([http://www.geneontology.org](http://www.geneontology.org)) and KOG
(http://www.ncbi.nlm.nih.gov/COG/) databases. Transcripts for which GO terms had been successfully mapped were then further annotated by linking the GO terms to associated functions according to the following parameters: (E-Value-Hit-filter 1.0E-6, Annotation cutoff 55, GO weight 5, Hsp-hit coverage cutoff 0).

4.4 Results

4.4.1 RNA Isolation and Assessment of Quality

The 4 sampling times and two strains for each sample provided 8 bulk RNA samples with total RNA yields ranging from 63 to 1655 μg with OD\textsubscript{260/280} ratios ranging from 1.56 to 2.12 (Table 4.1).

**Table 4.1** Total RNA extracted from *E. maxima* GS and *E. maxima* M6 from 3 morphological stages during sporulation and excysting sporozoites.

<table>
<thead>
<tr>
<th>Sample (sporulation time)</th>
<th>Volume (µl)</th>
<th>RNA concentration (ng/µl)</th>
<th>OD\textsubscript{260/280}</th>
<th>OD\textsubscript{260/230}</th>
<th>Total RNA (µg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS 18 h</td>
<td>100</td>
<td>1137.8</td>
<td>1.56</td>
<td>0.60</td>
<td>113.8</td>
</tr>
<tr>
<td>GS 22 h</td>
<td>100</td>
<td>16554.4</td>
<td>2.05</td>
<td>1.87</td>
<td>1655.50</td>
</tr>
<tr>
<td>GS 38 h</td>
<td>100</td>
<td>1559.8</td>
<td>1.94</td>
<td>0.71</td>
<td>156.0</td>
</tr>
<tr>
<td>GS sporozoites</td>
<td>100</td>
<td>629.2</td>
<td>2.00</td>
<td>0.95</td>
<td>63.0</td>
</tr>
<tr>
<td>M6 18 h</td>
<td>100</td>
<td>12198.3</td>
<td>2.15</td>
<td>2.36</td>
<td>1220</td>
</tr>
<tr>
<td>M6 22 h</td>
<td>100</td>
<td>9024.8</td>
<td>2.12</td>
<td>1.83</td>
<td>902.50</td>
</tr>
<tr>
<td>M6 38 h</td>
<td>100</td>
<td>2030.3</td>
<td>2.00</td>
<td>0.65</td>
<td>203.0</td>
</tr>
<tr>
<td>M6 sporozoites</td>
<td>100</td>
<td>3534.6</td>
<td>1.75</td>
<td>0.50</td>
<td>353.50</td>
</tr>
</tbody>
</table>

The isolated bulk cellular RNA was further purified to improve the quality and to enrich the mRNA by removing small RNA molecules from the sample as explained above. The RNA quality was then tested using formaldehyde gel. The ribosomal RNA bands (18S and 28S rRNA’s) were clear in gel (Figure 4.1). The appearance of the two bands show that the RNA was not degraded. The RNA samples were then used for ds cDNA synthesis.
Figure 4.1 RNA integrity was assessed using agarose formaldehyde gel electrophoresis of 1-3 µg of isolated bulk RNA. In all electrophoresis runs, bands corresponding to 18S and 28S rRNA were clearly visible. A. Lane 1 is marker and Lane 5 is RNA extracted from 22 h sporulation time point of *E. maxima* GS strain. B. Another sample of total RNA extracted from *E. maxima* M6.

A 647 bp fragment of the thrombospondin-related adhesive protein (TRAP – EmTFP250 [AY239227]) was amplified from the resulting ds cDNA to test the quality of the isolated RNA. This 647 bp fragment of the EmTFP250 gene was amplified successfully (Figure 4.2) indicating that the prepared cDNA was of good quality and was suitable for further processing and next generation sequencing.

Figure 4.2 A portion of the thrombospondin-related adhesive protein (‘TRAP’, EmTFP250 – see AY239227) was amplified from M6 to test the quality of the prepared cDNA and confirm its suitability for further processing and transcriptome generation using Illumina sequencing.
4.4.2 Illumina (NextGen) Sequencing

A total of $190 \times 10^6$ paired reads were generated from the *E. maxima* M6 cDNA library representing $\sim 38 \times 10^9$ bp of sequence data. A further $54 \times 10^6$ paired reads were generated from the *E. maxima* GS cDNA library totaling $\sim 10.8 \times 10^9$ bp of sequence data. All sequence data was exported from the Illumina sequencing platform as FASTQ format (Lipman and Pearson 1985; Cock et al. 2010) that incorporates the primary sequencing result as well as a quality assessment for each called base for each read. The base calls and their associated qualities are used for subsequent *de novo* assembly and further analyses.

4.4.3 Transcript Generation and Annotation from *E. maxima* GS and M6

*De novo* assembly of the paired-reads was completed independently for each *E. maxima* strain and involved generation of contig assemblies of the paired reads followed by generation of quality based consensus sequences from the assemblies to generate putative transcripts for each parasite strain. Hereafter, such consensus sequences will be referred to as ‘transcripts’.

For each strain, the *de novo* assembly procedure retained 10,000 assemblies of pair-reads from which the first 5,000 contigs based on read density (see section 4.3.6) from each strain were analyzed further; ordering of assemblies by Geneious is based on the total number of paired reads contributing to each assembly and transcript length for assemblies generated from the same number of sequencing reads. Top 5,000 contigs were selected based on number of reads contributing to each contig. For example, for *E. maxima* GS, assembly contig 1 was 3,746 bp long and consisted of 208,819 sequencing reads assembled together whereas assembly contig 4927 was assembled from 131 sequencing reads and was 1,058 bp in length (Figure 4.3). For each of the first 5,000 contig assemblies from each strain, a consensus sequence was generated from the aligned reads in the assembly based on quality of the contributing reads and assigning a quality score for the final consensus sequence.

The length of transcripts (cDNA’s) generated from *E. maxima* GS and *E. maxima* M6 using *de novo* assembly of paired-reads were similar between the two investigated strains. The transcripts obtained in the present study ranged from 164 to 8,670 bp in length.
for *E. maxima* GS and 111 to 8889 bp for M6 with an average length of 1597 and 1292 bp for *E. maxima* GS and M6, respectively. The majority of transcripts had lengths between 1000-2000 bp (Figure 4.3).

![Transcript length distribution](image)

**Figure 4.3** The distribution of transcript lengths of *E. maxima* GS and *E. maxima* M6. The blue columns represent GS strain. The longest transcript in GS strain was 8,670 bp. The red columns represent M6 strain. The longest transcript was 8,889 bp in this strain. The average length of transcripts of GS and M6 was 1,597 bp and 1,292 bp, respectively.

The resulting 5,000 transcripts from each strain were then searched against GenBank using Blast2GO to identify similar sequences and then gene orthology (GO) terms were mapped to each, where possible. For *E. maxima* GS, 2,894 of the 5,000 transcripts generated hits at a Blast Expect Value cutoff of 1.0E⁻³ when searched against the GenBank non-redundant nucleotide database (nr). Blast2GO mapped GO terms to 1,129 transcripts of these 2,894 hits (Figure 4.4). The majority of these transcripts (about 60%) was assigned GO terms associated with metabolic function and cellular processes (Table 4.2). Substantial numbers of transcripts were associated with a range of cellular functions such as biological regulation (7%), localization (7%), cellular component organization or biogenesis (4%) or developmental process (3%). Smaller numbers of transcripts were associated with additional GO terms (Figure 4.5).
Figure 4.4 The data distribution of 10,000 transcripts generated from *E. maxima* GS and *E. maxima* M6 blasted against NCBI using Blast2GO. The red bars represent M6 strain. Out of 5,000 transcripts of M6 strain, a total of 2569 transcripts had blast hits and distributed as follow: 955 transcripts were annotated, 753 transcripts had mapping results, and 861 transcripts had blast results only. The blue bars represent data obtained from GS strain. Out of 5,000 transcripts from GS, 2,893 transcripts had blast last hits and distributed as follow: 1,235 transcripts were annotated, 814 transcripts had mapping results, and 844 transcripts had blast results only. The total number of GS strain and M6 strain transcripts did not have blast hits against NCBI using Blast2GO were 2607 and 2431 respectively.

Of the top 5,000 transcripts generated by the *de novo* assembly in Geneious for *E. maxima* M6, 2,569 generated hits at a Blast ExpectValue cutoff of $1.0 \times 10^{-3}$ when searched against the GenBank non-redundant nucleotide database (nr). Blast2GO was able to map GO terms to 1,708 contigs of the 2,569 hits and further annotate 955 of the contigs for which GO terms were assigned. Of the 1,708 transcripts assigned GO terms by Blast2GO, the majority of these transcripts (~ 60%) were associated with metabolic function and cellular processes (Table 4.2). Similar to the *E. maxima* GS results, substantial numbers of transcripts were associated with a range of cellular functions such as biological regulation (7%), localization (7%), cellular component organization or biogenesis (4%) or developmental process (3%). Smaller numbers of transcripts were associated with additional GO terms (Figure 4.6).
Table 4.2 Percentages of transcripts for which putative biological function was assigned after blasting 10,000 transcripts generated from of *E. maxima* GS and *E. maxima* M6 against NCBI using Blast2GO. The highest percentage of transcripts (60%) among GS and M6 strains were assigned GO terms associated with metabolic and cellular process.

<table>
<thead>
<tr>
<th>Putative biological function</th>
<th>Percentage of transcripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metabolic process</td>
<td>30%</td>
</tr>
<tr>
<td>Cellular process</td>
<td>30%</td>
</tr>
<tr>
<td>Localization</td>
<td>6%</td>
</tr>
<tr>
<td>Biological regulation</td>
<td>6%</td>
</tr>
<tr>
<td>Response to stimulus</td>
<td>6%</td>
</tr>
<tr>
<td>Cellular component organization or biogenesis</td>
<td>4%</td>
</tr>
<tr>
<td>Organismal process</td>
<td>5%</td>
</tr>
<tr>
<td>Developmental process</td>
<td>3%</td>
</tr>
<tr>
<td>Signaling</td>
<td>2%</td>
</tr>
<tr>
<td>Reproduction</td>
<td>2%</td>
</tr>
<tr>
<td>Surface antigen</td>
<td>1%</td>
</tr>
<tr>
<td>Death</td>
<td>1%</td>
</tr>
<tr>
<td>Growth</td>
<td>1%</td>
</tr>
<tr>
<td>Locomotion</td>
<td>1%</td>
</tr>
</tbody>
</table>

Large numbers of transcripts from both strains of *E. maxima* had hits in GenBank but no GO terms were assigned (861 contigs from M6 and 1765 contigs from M6). These transcripts had no significant matches to proteins to which GO terms had been assigned; therefore, it is possible that these represent novel genes or apicomplexan genes for which GO terms have not been defined within public databases. Although the genome of *E. maxima* (Weybridge strain) has been shotgun sequenced (Reid et al 2013, personal communication), and is publicly available, annotation of these contigs has not yet been made public; this makes mapping and annotation of transcripts challenging. However, the *Eimeria* hits are previous EST and cDNA sequences that have been uploaded to GenBank by other researchers previously.
Figure 4.5 Biological process distribution of 2,067 transcripts (of 5,000 total transcripts examined in detail) generated from *E. maxima* GS for which GO terms could be assigned using Blast2GO (http://www.blast2go.com/b2ghome) and public sequence databases. The majority of these transcripts (1,253 transcripts) were associated with GO terms for metabolic and cellular process. Transcripts that were identified as putative surface antigens (SAgs) numbered 19 transcripts.
Figure 4.6 Biological process distribution of 1,610 transcripts (of 5,000 total transcripts examined in detail) generated from *E. maxima* M6 for which GO terms could be assigned using Blast2GO (http://www.blast2go.com/b2ghome) and public sequence databases. The majority of these transcripts (970 transcripts) were associated with GO terms for metabolic and cellular process. Transcripts that were identified as putative surface antigens (SAgs) numbered 18 transcripts.
The *E. maxima* GS and *E. maxima* M6 transcripts showed similarities with sequences from other parasites, particularly from the phylum Apicomplexa. Of transcripts from both strains (GS and M6), the largest number had homology with genes identified in *Toxoplasma gondii*. Of 2,894 transcripts of GS strain that had hits following a Blast search, 1,147 transcripts (39.6%) hit *Toxoplasma gondii*; likewise, 985 of 2,569 transcripts (38.3%) of the M6 strain had Blast hits with *T. gondii* because of the availability of a robust *T. gondii* genome that is well annotated (Figure 4.7).

![Figure 4.7](http://www.blast2go.com)

**Figure 4.7** Top hits species distribution of 10,000 transcripts generated from *E. maxima* M6 and *E. maxima* GS strains against the nr database of NCBI using Blast2GO. The red columns represent GS strain. The total number of GS transcripts that hit matches in other species was 2893. The majority of transcript hits matched sequences from *Toxoplasma gondii*. The blue column represents the transcripts of M6 strain that had Blast hits with parasite sequences. The total number of M6 transcripts that had Blast hits was 2,532. The majority of these transcripts matched sequences from *Toxoplasma gondii*.

One group of transcripts for which BLAST hits were found with sequences in GenBank but that did not have GO terms assigned were found to encode putative surface antigens. There were 19 SAgs identified among the contigs generated from *E. maxima* GS and a further 18 SAgs identified from *E. maxima* M6. Many of these transcripts encoded proteins that shared similarities to surface antigen superfamily cl12617, a eukaryotic
family of surface antigens that have shared structures. The majority of these SAgs were expressed by both *E. maxima* GS and M6 (31 SAgs representing 15 genetic loci) and the remaining 6 SAgs represented uniquely expressed genetic loci in the two strains. Identifying and comparing these SAgs from the GS and M6 strains of *E. maxima* and their relative expression during sporulation and excystation is addressed in chapter 5 of this thesis.

Searching the top 5,000 transcripts generated from *E. maxima* GS against the top 5,000 transcripts generated from *E. maxima* M6 using a custom Blast search indicated that there were approximately 4,600 transcripts that shared sequence similarities between the strains (4,578 or 4,740 depending on the direction of the Blast search). Of the approximately 400 transcripts that were not shared between the two strains (360 or 422 depending on the direction of the Blast search), nearly half (167 or 192, respectively) were found to be shared when the Blast search was expanded to include all 10,000 transcripts generated during de novo sequence assemblies from each strain. This suggests that at least some of the apparent differences between these strains actually reflect differences of expression rather than differences in the genes themselves.

### 4.5 Discussion

Next generation sequencing has provided a means of accessing the existence and primary structure of transcribed genes from a single species, stage of a species or even a short developmental period during the life cycle of a species. The biological function of such genes can then be predicted efficiently (although not easily) with the automatic functional annotation servers that are freely available, such as Blast2GO ([http://www.blast2go.com/b2ghome](http://www.blast2go.com/b2ghome)). The understanding of functional annotation of a large number of genes allows us to categorize them according to their function to classes and subclasses such metabolic process, immune system process, locomotion, biological adhesion, death, antigenicity and so on. In the present study, cDNA samples were prepared from mRNA extracted from three stages during sporogonic development (18 h, 22 h, and 38 h of sporulation, see Al-Badri and Barta 2012) as well as excysting sporozoites so that the resulting transcriptomes from these two strains covered the entire sporulation and excystation process. This study was conducted to identify the biological function of large
number of transcripts (10,000 in total from the GS and M6 strains); these data were then analyzed further as part of a comprehensive study to compare two immunologically distinct strains of *E. maxima* (GS and M6).

As an indication of the quality of the transcripts obtained in the present study, one of the longer transcripts from *E. maxima* M6 was analyzed in some detail. Assembly contig 62 of *E. maxima* M6 generated a transcript 8,353 bp long based on 18,712 reads and was found to code for a highly conserved member of the TRAP protein family (EmTFP250) for which a complete mRNA sequence is available (GenBank AY239227.1). This protein was identified in *E. maxima* Houghton strain. The homologous transcript from *E. maxima* GS was 8,440 bp long based on 6,588 reads and 100% pairwise identical to *E. maxima* M6. Pairwise comparison of either the M6 or GS transcript with the EmTFP250 mRNA sequence showed > 99.9% sequence identity (5 SNP’s identified over the length of the transcript; 3 within the CDS, 2 non-synonymous). These observations and previously noted confirmation that the cDNA could be generated and amplified using EmTFP250-specific primers (see above) suggest that cDNA library construction, subsequent sequencing and *de novo* assembly were able to generate high quality transcripts from these *E. maxima* strains, including mRNA’s that exceeded 8 kb in length.

The transcripts obtained in the present study ranged from 164 to 8,670 bp in length for *E. maxima* GS and 111 to 8,889 bp for *E. maxima* M6 with an average length of 1597 and 1292 bp for each strain, respectively. These numbers are in rough agreement with cDNA-based transcriptome analyses of other *Eimeria* species where full-length cDNA sequences ranged from 441 to 3,083 bp in length and averaged 1,647 bp (Amiruddin et al. 2012). However, in the present study, large numbers of longer transcripts were obtained compared with those reported for cDNA-based transcriptomes analyses; for example, 248 transcripts exceeded 3.5 kb for the *E. maxima* GS *de novo* assembly and 162 transcripts from *E. maxima* M6 were longer than 3.5 kb. The notably longer maximum lengths of transcripts identified in the present NGS-based transcriptome analyses likely reflect limitations in the generation of long cDNA clones rather than an actual difference in the expression of longer transcripts in *E. maxima* GS and M6 compared with the other sampled species and stages.
In the present work, transcriptomes were generated from the sporulation and excystation of sporozoites of *E. maxima* for the first time; other life cycles stages of *E. maxima* and of other *Eimeria* species have been examined previously by others. For example, Schwarz et al. (2010) examined the transcriptome of merozoites of *E. maxima* Tysons strain and compared the transcriptomes of the latter species to those of *E. tenella* and *E. acervulina* using EST analysis. Out of 2,680 individual EST’s obtained from merozoites of *E. maxima*, only 48.2% had significant (E < 10^{-5}) similarities to sequences with other parasites (primarily *Eimeria* spp. or *Toxoplasma gondii*), and 47.5% were unique (Schwarz et al. 2010). Our transcripts also showed similarities with other protozoan parasites. Out of 5,000 transcripts of GS strain, 2,898 transcripts showed that they have similarities with other alveolate protozoa. Similarly, 2,569 transcripts out of 5,000 contigs of M6 strain showed similarities with other protozoan parasites (Figure 4.7). The majority of the transcripts of GS and M6 strains hit *T. gondii* sequences, likely because of the large amount of sequence data that is available for this parasite (Sibley and Boothroyd 1992; Kissinger et al. 2003) and the comparative maturity of the annotation for these sequences.

Schwarz et al. (2010) annotation of EST’s with Blast hits from *E. maxima* merogonic stages showed that most were associated with cell functions such as translation, cytoskeleton, metabolism, signaling, transport, and protein folding but a small number were associated with surface antigens. For the GS strain, the biological functions of 2,073 transcripts were determined (Figure 4.1). The largest component of annotated transcripts (30%) mapped to GO terms associated with cellular or metabolic functions. Similar results were obtained from *E. maxima* M6, where 1,592 of 5,000 transcripts could be associated with a putative function (Figure 4.2). Interestingly, the proportion of annotated transcripts from M6 associated with metabolic functions (31%) was essentially the same as for *E. maxima* GS.

A number of putative surface antigens (SAgs) were identified (10 SAg clusters comprised of 63 ESTs) among the 416 total annotated EST assemblies generated by Schwarz et al. (2010). In a similar study, complete coding sequences identified 25 putative secretory and 60 putative surface proteins after the cDNA sequences from *E. tenella* second generation merozoites were obtained (Amiruddin et al. 2012). Surface antigens
(SAgs) were identified among the 10,000 transcripts that were analysed from *E. maxima* GS and M6 sporozoite transcriptomes in the present study. The 5,000 GS transcripts had 19 putative surface proteins that hit paralogous sequences of known surface antigens. A similar number of putative surface proteins (18) that hit paralogous sequences of known surface antigens were detected among the 5,000 transcripts of M6 strain. Identifying and characterizing these SAgs between these two parasite strains may reveal differences in their surface antigens. Any such results of comparing SAgs of these two strains might explain the lack of cross immunity between them which may eventually lead to developing a subunit vaccine (Beattie et al., 2001). At the very least, further characterization and identification of SAgs that are expressed by both strains and that are identical, or nearly so, would suggest that such SAgs are unlikely to contribute to strain specificity of the immune response. The nature of the SAg repertoire of each *E. maxima* strain is explored in more detail in Chapter 5.

The proteome of different stages of life cycle (unsporulated oocysts, sporulated oocysts, sporozoites, and second-generation merozoite) among different *E. tenella* life cycle was investigated by Lal et al. (2009). There were 68 GPI-anchored surface proteins identified in the merozoites of this parasite; but the number of GPI-anchored proteins identified in the sporozoite stage was 16 SAgs. In comparison to our study, the numbers of putative SAgs transcribed at the mRNA level during sporogony of *E. maxima* GS and *E. maxima* M6 stage were 19 and 18, respectively. The results of our study and other studies suggested that more numerous SAgs may be produced during merogonic development than during sporogony, including the production of sporozoites, of *Eimeria* species.

Many of transcripts for which BLAST hits were found did not have GO terms because publically available information concerning the genome of *E. maxima* is still minimal. The genome sequence of this coccidian parasite has not yet been released which makes identifying, mapping and annotating newly generated transcripts difficult and challenging. Other parasites, such as *Toxoplasma gondii*, have a well-annotated public genome that has been investigated more thoroughly. The investigation of EST’s, cDNA or transcriptome sequences from various *Eimeria* spp. clearly demonstrates that there remain large numbers of proteins in apicomplexan parasites that have not been annotated and for
which the function remains unknown. For example, a variety of workers found that 47% to 73% of putative protein sequences with BLAST hits in the GenBank database were lacking annotation(s) and described as ‘unknown’ (Wan et al.; 1999; Ng et al. 2002; Miska et al.; 2004; Schwarz et al. 2010).

Novas et al. (2012) observed that 90% of the cDNA’s generated from different stages (unsporulated oocysts, sporoblastic oocysts, sporulated oocysts, sporozoites and second generation merozoites) of *Eimeria* spp. encoded a putative protein coding sequence. However, only a small subset (32–38%) of these hypothetical proteins matched sequences in the non-redundant (nr) database of GenBank following BLASTx searches. Aarthi et al. (2011) studied the expressed sequence tags generated from sporozoites of *Eimeria brunetti*. In the latter study, 283 ESTs were generated from a sporozoite cDNA library but only 50 ESTs had matches to proteins in NCBI’s non-redundant protein database using BLASTx (E-value cutoff ≤ 1e−05). Whereas, the remaining 233 transcripts had no significant matches with known proteins and therefore were interpreted to represent novel genes. The number of unknown transcripts (without GO terms) in our data is not unexpected because 50% of EST’s from other well characterized apicomplexan parasites remain unknown to date (Li et al. 2003; Cui et al. 2005; Dybas et al. 2008). Of the large number of transcripts without GO terms identified in the present study, most are likely to represent well known (or at least previously reported) proteins that lack any detailed annotations and a minority may represent novel proteins.

In conclusion, the transcriptomes from *E. maxima* GS and *E. maxima* M6 during sporulation and sporozoite excystation were generated for the first time in this study. The analysis of transcriptomic sequences (transcripts) led to the identification of a number of genes that likely code for GPI-anchored SAgS that were expressed by both strains (31 SAgS representing 15 genetic loci); an additional 6 SAg transcripts encode apparently uniquely expressed SAgS in the two strains. The identified putative surface antigens will be comprehensively studied in an effort to explain the lack of cross-immunity between these two closely related strains of *E. maxima*. Understanding the differences between these parasites may help to identify molecules that contribute to the lack of immunological cross-reactivity of the GS and M6 strains and this may ultimately help in the development of a protective subunit vaccine.
Chapter 5- Identification of glycophosphatidylinositol (GPI) anchored surface proteins and comparison between two immunologically distinct strains of *Eimeria maxima* (GS and M6)

5.1 Abstract

The glycophosphatidylinositol (GPI) anchored surface proteins expressed by sporozoites of two immunologically distinct strains of *Eimeria maxima* GS and M6 were investigated in this study. The transcriptome of sporozoites of these two strains were sequenced using Illumina after mRNA samples were obtained from partially sporulated oocysts (18 h, 22 h, and 38 h of sporulation) and fully excysting sporozoites. Eighteen transcripts encoding putative glycophosphatidylinositol (GPI) anchored proteins were identified from each strain. Most of these putative surface antigen transcripts were well conserved between the two strains (15 SAg genes expressed by both strains) but 3 GPI-anchored SAg transcripts (3 SAg genes) were expressed within the transcriptome of the GS strain only; likewise, 3 GPI-anchored SAg transcripts (3 SAg genes) were expressed in the transcriptome of only the M6 strain. Polymorphism of orthologs was identified among shared SAgs from GS, M6, Wey, and H strains with \( \delta_N/\delta_S \) ratios ranging from 0 to 2.37. A large multicopy family of surface antigens (at least 9 tandemly arrayed gene copies with repetitive structures favouring recombination) was identified that showed similarity with families of similar surface antigens in other apicomplexan parasites. Interestingly, in a second multicopy gene family comprised apparently of only three gene copies, 2 of 3 of these genes were expressed by the GS strain alone and the remaining gene was expressed by M6 only; comparative genome sequences indicated that all three genes likely exist in both *E. maxima* GS and M6. The selective expression of 1 or 3 genes as observed in the present transcriptome analysis may contribute to the lack of cross immunity between these parasites.

5.2 Introduction

*Eimeria maxima* is an economically important parasite that impacts the health of chickens globally. This species is highly immunogenic and infection with as few as five oocysts is capable of inducing protective immunity against homologous challenge (Blake et al. 2005). A comparative study demonstrated that *E. maxima* is the most immunogenic species of avian coccidia (Rose and Long 1962) as later confirmed by other researchers (Joyner and Norton 1976; Long and Millard 1977). Immunogenicity of *E. maxima* is so strong that partial immunity against *E. maxima* infection could be achieved by inoculation of only a single sporocyst (equivalent to 2 individual sporozoites) to chickens (Lee and Fernando 1978).

Two immunologically distinct strains, GS and M6, of *Eimeria maxima* (see Beattie et al. 2001) were investigated in this study. Despite their relative lack of immunological cross-protection, these two strains are similar in their kinetics of oocyst sporulation and
oocyst production (Al-Badri and Barta 2012). The Guelph strain of *Eimeria maxima* (*E. maxima* GS) was isolated from an Ontario, Canada, poultry facility in 1973 and a single-oocyst derived line has been maintained at the University of Guelph since then. *Eimeria maxima* M6 was derived from parasites initially obtained on a Florida, USA, broiler farm in 1994. Each of these two strains has been characterized extensively (Martin et al. 1997; Barta et al. 1998; Beattie et al. 2001; Ogedengbe et al. 2011; Al-Badri and Barta 2012).

Antigenic diversity is a trait and defense mechanism of some pathogenic agents such as *Eimeria maxima* (see Fitz-Coy 1999; Smith et al. 2002). There are two known classes of antigenic diversity. The first class in which the infection with one subpopulation may fail to protect against subsequent infection results from the genetic diversity within the pathogen population (Fitz-Coy 1992; Long 1974; Tomley 1994). However, the second class of antigenic diversity comes as a result of selective expression of members of large polymorphic gene families coding for surface proteins such as exploited by *Trypanosoma* and *Plasmodium* species by individual parasites. In addition to the selective expression of genes in multicycopy genetic loci, the repetitive structure of such loci may promote recombination among such gene copies, further enhancing genetic diversity. Pathogens that have one (or both) of these mechanisms can prolong their persistence in their host because they can present antigens that avoid an established immune response either through lack of expression or antigenic polymorphism (antigenic drift). For these reasons, vaccines are difficult to develop or fail to protect all hosts infected with these types of pathogens (Borst et al. 1997; Deitsch et al. 1999; Newbold 1999). Despite extensive efforts over decades to develop a vaccine against the virulent human malarial parasite, *Plasmodium falciparum*, such antigenic variation at the individual parasite and at the population level has thwarted development of an effective vaccine (Birkett et al. 2013).

Glycophosphatidylinositol (GPI) anchored proteins (antigens) are expressed on the surface of apicomplexan parasites such as *Toxoplasma gondii*, *Sarcocystis neurona*, *Plasmodium falciparum*, and *Eimeria* species. These surface antigens (SAgs) are well known to elicit host immune responses and also have important roles in parasite biology affecting pathogenicity through their involvement in host cell invasion and immune evasion (Taberes et al. 2004; Jung et al. 2004; Gilson et al. 2006). Schofield and Hackett found that GPI-anchored proteins obtained from the erythrocytic stage of *P. falciparum*
were able to stimulate the synthesis of IL-1 and tumor necrosis factor-α (TNF-α) by murine macrophages. Moreover, Tachado et al (1996) illustrated that malarial GPI-anchored proteins elicited de novo expression of nitric oxide synthase indicating macrophage stimulation.

Coccidiosis in chickens, caused by seven species of *Eimeria*, is still controlled mainly by adding chemical and/or ionophorous anticoccidials to chicken feed prophylactically or through use of live coccidiosis vaccines (live attenuated or virulent) administered at the hatchery. In-feed antimicrobials are not desirable from both drug resistance and public acceptance perspectives, and live vaccines are comparatively expensive with risks of vaccine-associated negative impact on bird health. In particular, live vaccines cause sufficient alteration to the intestinal tract of vaccinated birds (especially elevated mucus production) that such birds are predisposed to necrotic enteritis caused by *Clostridium perfringens* (see Pedersen et al. 2008; Park et al. 2008; Timbermont et al. 2011). An effective subunit vaccine might offer a less costly and safer alternative for controlling coccidiosis. The only commercially marketed subunit vaccine against coccidiosis, CoxAbic®, (Phibro Animal Health Corporation, Teaneck, NJ) was based on soybean lectin affinity-purified antigens of *E. maxima* macrogamonts. This vaccine is used by inoculation into broiler breeders; the resulting broiler offspring demonstrate limited resistance to coccidial challenge for the first few weeks after hatch (Sharman et al. 2010), apparently through transfer of maternal antibodies through the yolk. However, this resistance to infection is far less protective than the sterile immunity conferred by a natural infection or live vaccines.

The objective of this study was to identify and compare the transcribed SAgS from *E. maxima* GS and *E. maxima* M6 and study the abundance of their expression in the transcriptomes of these parasites. Any identified variant SAg transcripts, or differences in expression, could explain the lack of cross immunity between these two strains. Such strain-specific SAGs would be rational targets for the development of a subunit vaccine to protect chickens against infection with *E. maxima* during the early stages of infection and before damage occurs in the intestine.
5.3 Materials and Methods

5.3.1 Complementary DNA construction

Samples of RNA were collected from *E. maxima* GS and *E. maxima* M6 from three different stages during sporulation of oocysts (18 h, 22 h and 38 h of sporulation at 26°C (Al-Badri and Barta 2012) and from mature sporozoites undergoing excystation. Isolated RNA from each sample (8 in total – see Table 4.1) was then column-purified to remove small RNA fragments. Detailed description of the RNA isolation, purification and cDNA library construction is provided in Chapter 4. Briefly, the purified bulk cellular RNA was pooled quantitatively from the four samples from each strain prior to selection of mRNA using oligo-dT columns and cDNA production. The cDNA from each strain was fragmented and joined to strain-specific adaptors before paired-read sequencing using an Illumina HiSeq 2000 instrument (see Chapter 4).

5.3.2 De novo assembly of paired reads

Initial *de novo* assembly of 100 bp paired reads generated by Illumina sequencing was accomplished from $8 \times 10^6$ reads (approximately 800 Mb of sequence data/strain) from each of *E. maxima* M6 and *E. maxima* GS using the *de novo* assembly function within Geneious Pro (version 6 and later, Drummond et al. 2012). The *de novo* assembly parameters were as follow: maximum ambiguity (4), index word length (10), maximum gap size (1), maximum mismatches per read (10%), and inter-read distances were used to improve assembly. For each *E. maxima* strain, 10,000 consensus sequences (conseqs) were retained at the conclusion of *de novo* assembly for further analyses. Preliminary identification and annotation of the contigs was accomplished using Blast2GO’s (Conesa et al. 2005) Blast, Mapping and Annotation functions ([www.blast2go.com](http://www.blast2go.com)) as described in detail in Chapter 4.

For each *E. maxima* strain, consensus sequences from each of the 10,000 contigs output from the *de novo* assembly were then used to generate a custom BLAST database within Geneious. Contigs previously identified as potential surface antigens using Blast2GO were then searched against this custom database within Geneious (using both blastn and blastx search algorithms) to assist in the identification of similar contigs from each parasite strain. In this way, as many contigs encoding potential surface and/or
GPI-anchored proteins as possible were identified among the 10,000 contigs generated for each strain. All such contigs, regardless of contig length, were retained for further analysis and characterization using prediction tools listed in materials and methods (5.3.5).

5.3.3 Mapping of transcripts against genomic sequences from *E. maxima* Houghton Strain and *E. maxima* Weybridge Strain

Our preliminary transcripts encoding potential surface antigens were then aligned against genomic sequences of *E. maxima* Houghton strain that is available freely (http://www.genomemalaysia.gov.my/emaxdb/) and *E. maxima* Weybridge strain (ftp://ftp.sanger.ac.uk/pub/pathogens/ar11/assemblies/). For transcripts that were relatively conserved in primary sequence, unambiguous assignment of such transcripts from M6 and GS could be made readily. In these cases, the gene nomenclature of Reid et al. (2013, personal communication) was assigned to these relatively conserved transcripts. In addition, the UTR regions, exons and introns were annotated for each of the reference genomic sequences. Thus, for many transcripts generated by *de novo* assembly of the paired-read sequence data, UTR and coding sequences (CDS) for four strains of *E. maxima* were available for further study.

However, it became clear that some contigs could only be assigned tentatively to one of several closely related members of one or more multicopy gene families. Examination of the paired-read sequence assemblies that generated the transcript sequences revealed that multiple, distinct transcripts had been combined into an artificially reduced number of contigs that contained a few to many nucleotide positions at which competing base assignments had been made in the consensus sequence. Nonetheless, these “chimeric” conses could be aligned unambiguously with one or more positions on the genomic sequences of Houghton strain, Weybridge strain or both. For the contigs of GS or M6 that clearly contained multiple transcripts, exons were annotated on our transcripts by comparison with either the h or Wey genome reference alignment. For these chimeric conses expressed by GS or M6 aligning with multicopy surface antigen genes of H or Wey, there were 4 exons identified and each of these was individually extracted from our conses. The extracted exons were then used as reference sequences to assemble *E. maxima* M6 and *E. maxima* GS paired reads using highly discriminating parameters in
order to avoid ambiguous assemblies or assemblies that combined reads from several related but distinct templates. The resulting assemblies were examined and consensus sequence of each assembly was generated. Each validated exon conseq included 3’ and 5’ tails extending beyond the reference sequences and thus these validated exons could then be assembled back into conseqs covering the entire CDS for SAg transcripts. To validate the existence of these potential transcripts in the transcriptomes of GS and M6, each new conseq was used as a reference sequence for assembly of all paired-reads from M6 and GS; the reference assembly parameters used were highly restrictive so that only closely mapped paired-reads with 1% or less mismatches or ambiguities were included. Each resulting sequence assembly was examined for directionality of, and distances between, mapped paired-reads to confirm that each assembly represented a valid contiguous transcript. All potential SAg transcripts were validated in this manner prior to use in further analyses.

5.3.4 Relative abundance of gene (Transcript) expression

An estimate of the relative abundance of each validated transcript within the transcriptomes of each *E. maxima* strain was determined by using each transcript sequence as the reference for 2 highly discriminatory reference assemblies within Geneious using 376 million paired reads of *E. maxima* M6 and 104 million paired reads from *E. maxima* GS, i.e. all available Illumina sequence data. Reference assembly parameters permitting 1% or less mismatches or ambiguities and the assembly restriction that only pair-reads that mapped nearby be used for the assemblies ensured that these assemblies were highly discriminatory. The number of paired-reads from GS or M6 that assembled to each reference sequence was recorded over the entire transcript as well as the CDS for each transcript. Relative abundance at the transcript level was assumed to be proportional to the relative number of paired reads that assembled with high fidelity to each validated transcript. Longer transcripts would be expected to have more reads mapping to them simply because of their lengths; therefore, a transcript length normalized relative abundance was calculated using the equation below:

\[
\text{Transcript Relative Abundance} = \frac{\text{Number of reads mapped to Transcript} \times 100000000 \text{ [Constant]}}{\text{Total number of paired reads} \times \text{Transcript Length}}
\]
The same equation was used to determine the CDS length normalized relative abundance by substituting the number of reads mapped to the CDS and the CDS length for numbers associated with the entire transcript.

\[
CDS\ Relative\ Abundance = \frac{\text{Number of reads mapped to CDS} \times 100000000 [\text{Constant}]}{\text{Total number of paired reads} \times \text{CDS Length}}
\]

Once the relative abundance for all transcripts was determined, the relative abundance of each transcript or CDS was transformed so that the least abundant transcript or CDS was assigned a relative abundance of 1.

If a particular genetic locus appeared to be lacking from either the GS or M6 transcriptome, a reference sequence assembly was attempted using all comparative sequences as reference sequences and all paired-reads from the \textit{E. maxima} strain that did not appear to express the gene in question. For example, for the Type I SAgs described in detail below, Type I.3 is expressed by \textit{E. maxima} GS and the gene is found on both \textit{E. maxima} Wey and H strains at near identity. To confirm a lack of expression of this genetic locus by \textit{E. maxima} M6, the mRNA transcript of GS as well as the genome based sequences for the mRNA for this product from \textit{E. maxima} Wey and \textit{E. maxima} H, were each individually used as the reference sequence for assembly of all available Illumina sequence reads; a lack of a contiguous assembly across the CDS from all available reference sequence assemblies was used as evidence that a particular genetic locus was represented or not in the transcriptome from a particular \textit{E. maxima} strain.

5.3.5 \textit{Surface antigens (SAgs) characterization}

Transcripts from both strains that had similarity to surface antigens of \textit{Eimeria} spp., as identified above, were examined for open reading frames to determine potential translated amino acid sequences. The coding region of each contig was translated. The translations of proteins were studied to find signal peptides, glycophosphatidylinositol cleavage sites (GPI), and transmembrane structures using several predicting algorithms. Signal peptide presence was predicted using Signal-3L predicting server (http://www.csbio.sjtu.edu.cn/bioinf/Signal\_3L/) available freely online and SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/) server (Shen and Chou 2007). Potential GPI
cleavage sites were predicted using big PI predictor server available online (http://mendel.imp.ac.at/gpi/gpi_server.html) as well as the publically accessible server PredGPI (http://gpcr2.biocomp.unibo.it/predgpi/pred.htm), (Pierleoni et al. 2008). The transmembrane structure of surface proteins was predicted using a Hidden Markov Model (HMM) within Geneious. Protein secondary structure was predicted using the emboss garnier program implemented from within Geneious; the Garnier program uses the Garnier Osguthorpe Robson algorithm (GOR I) for secondary structure predictions (Garnier et al. 1978).

5.4 Results

A total of 190×10^6 paired reads were generated from *E. maxima* M6 representing ~38×10^9 bp of sequence data. A further 54×10^6 paired reads were generated from *E. maxima* GS total ~10.8×10^9 bp of sequence data. Paired-reads (200 bp total sequence data with an intervening region of approximately 50 unknown nucleotides) were generated by connecting associated single reads (each representing 100 bp sequence data) within Geneious.

There were 18 distinct transcripts (encoding GPI-anchored SAGs) identified from *E. maxima* GS and 18 distinct transcripts (encoding GPI-anchored SAGs) were discovered in *E. maxima* M6. Despite considerable differences in the primary sequences of each of these confirmed transcripts, each CDS encoded a peptide that had highly conserved features Figure 5.1.
Figure 5.1 Typical surface antigen gene, transcript and translation product. (A) Typical organization of a surface antigen gene (Example: Type B.1.M6). The blue bar represents the complete gene locus on the genome. The green bars represent 5’ (upstream) and 3’ (downstream) untranslated regions (UTR). There are four exons that represent the coding region of the gene (yellow). The dark grey bars are introns (non-coding sequences) which are nucleotide sequence within a gene that are removed by RNA splicing while the final mature RNA product of a gene is being generated (Deutsch and Long 1999). Introns start with bases GT and end with bases AG. (B) After introns have been removed by RNA splicing, the four exons (yellow color) remains present within the final mature mRNA product of that gene and forming the coding region (CDS) that will be translated into a protein (GPI anchored in this example). (C) Transmembrane structure of the translated CDS from a surface antigen transcript. SAg translation products showed 2 transmembrane structures at the N- and C-terminal regions of the propeptide (light blue). The N-terminal transmembrane structure corresponded to a signal peptide region (amino acids 1–23) predicted using Signal3L. The transmembrane structure at the C-terminus corresponded to the portion of the propeptide that was cleaved to attach the mature peptide (green) to the GPI anchor cleavage site (dark blue) that was predicted to be at amino acid 218 (S) according to big-PI predictor.
Table 5.1 Description of 36 transcripts identified in *E. maxima* GS and *E. maxima* M6. The transcripts were assigned to genetic locus identities by examining their similarity to named SAg genes that had been annotated as part of a nuclear genome sequencing effort on the *E. maxima* Weybridge strain (Reid et al., 2013 in prep.). Translation of *E. maxima* GS and *E. maxima* M6 CDS’s showed that the all putative proteins have signal peptide. The entire AA showed that they have (GPI) cleavage site using big-PI Predictor.

<table>
<thead>
<tr>
<th>Single or Multi-Copy Gene Family</th>
<th>Transcript Name</th>
<th>Gene Nomenclature (c.f. Weybridge)</th>
<th>Transcript Length (nt)</th>
<th>CDS length (nt)</th>
<th>Peptide Length (AA)</th>
<th>Signal peptide prediction (AA)</th>
<th>GPI-cleavage site (AA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multicopy gene family 1 (Type A SAGs)</td>
<td>Type A.1.GS</td>
<td>EMWEY_00001892</td>
<td>1204</td>
<td>1053</td>
<td>350</td>
<td>1-23</td>
<td>328 (G)</td>
</tr>
<tr>
<td></td>
<td>Type A.1.M6</td>
<td>EMWEY_00001892</td>
<td>1545</td>
<td>1053</td>
<td>350</td>
<td>1-23</td>
<td>328 (G)</td>
</tr>
<tr>
<td></td>
<td>Type A.2.GS</td>
<td>EMWEY_00058430</td>
<td>1560</td>
<td>1086</td>
<td>361</td>
<td>1-15</td>
<td>339 (G)</td>
</tr>
<tr>
<td></td>
<td>Type A.3.GS</td>
<td>EMWEY_00059060</td>
<td>1588</td>
<td>1026</td>
<td>341</td>
<td>1-15</td>
<td>319 (G)</td>
</tr>
<tr>
<td></td>
<td>Type A.3.M6</td>
<td>EMWEY_00059060</td>
<td>1343</td>
<td>1026</td>
<td>341</td>
<td>1-15</td>
<td>319 (G)</td>
</tr>
<tr>
<td></td>
<td>Type A.4.GS</td>
<td>EMWEY_00059060</td>
<td>1401</td>
<td>1050</td>
<td>349</td>
<td>1-15</td>
<td>326 (N)</td>
</tr>
<tr>
<td></td>
<td>Type A.4.M6</td>
<td>EMWEY_00059060</td>
<td>1173</td>
<td>1050</td>
<td>349</td>
<td>1-15</td>
<td>326 (N)</td>
</tr>
<tr>
<td></td>
<td>Type A.5.M6</td>
<td>EMWEY_00059740</td>
<td>1254</td>
<td>1089</td>
<td>362</td>
<td>1-15</td>
<td>340 (G)</td>
</tr>
<tr>
<td></td>
<td>Type A.6.GS</td>
<td>EMWEY_00052830</td>
<td>1471</td>
<td>1092</td>
<td>363</td>
<td>1-15</td>
<td>341 (G)</td>
</tr>
<tr>
<td></td>
<td>Type A.6.M6</td>
<td>EMWEY_00052830</td>
<td>1454</td>
<td>1092</td>
<td>363</td>
<td>1-15</td>
<td>341 (G)</td>
</tr>
<tr>
<td></td>
<td>Type A.7.M6</td>
<td>EMWEY_00052820</td>
<td>1315</td>
<td>1002</td>
<td>334</td>
<td>1-15</td>
<td>313 (S)</td>
</tr>
<tr>
<td></td>
<td>Type A.8.GS</td>
<td>EMWEY_00059740</td>
<td>1425</td>
<td>1101</td>
<td>366</td>
<td>1-15</td>
<td>344 (G)</td>
</tr>
<tr>
<td></td>
<td>Type A.8.M6</td>
<td>EMWEY_00059740</td>
<td>1354</td>
<td>1101</td>
<td>366</td>
<td>1-15</td>
<td>344 (G)</td>
</tr>
<tr>
<td>Multicopy gene family 2 (Type B SAGs)</td>
<td>Type B.1.GS</td>
<td>EMWEY_00005470</td>
<td>1733</td>
<td>714</td>
<td>237</td>
<td>1-23</td>
<td>218 (S)</td>
</tr>
<tr>
<td></td>
<td>Type B.1.M6</td>
<td>EMWEY_00005470</td>
<td>1610</td>
<td>714</td>
<td>237</td>
<td>1-23</td>
<td>218 (S)</td>
</tr>
<tr>
<td></td>
<td>Type B.2.GS</td>
<td>EMWEY_00005480</td>
<td>1288</td>
<td>726</td>
<td>241</td>
<td>1-23</td>
<td>222 (S)</td>
</tr>
<tr>
<td></td>
<td>Type B.2.M6</td>
<td>EMWEY_00005480</td>
<td>1338</td>
<td>726</td>
<td>241</td>
<td>1-23</td>
<td>222 (S)</td>
</tr>
<tr>
<td></td>
<td>Type B.4.GS</td>
<td>EMWEY_00005500</td>
<td>1144</td>
<td>726</td>
<td>241</td>
<td>1-23</td>
<td>222 (S)</td>
</tr>
<tr>
<td></td>
<td>Type B.4.M6</td>
<td>EMWEY_00005500</td>
<td>1284</td>
<td>726</td>
<td>241</td>
<td>1-23</td>
<td>222 (S)</td>
</tr>
<tr>
<td></td>
<td>Type B.5.GS</td>
<td>EMWEY_00005510</td>
<td>1464</td>
<td>876</td>
<td>291</td>
<td>1-25</td>
<td>269 (N)</td>
</tr>
<tr>
<td></td>
<td>Type B.5.M6</td>
<td>EMWEY_00005510</td>
<td>1366</td>
<td>876</td>
<td>291</td>
<td>1-25</td>
<td>269 (N)</td>
</tr>
<tr>
<td>Multicopy gene family 3 (Types C, D, G, J SAGs)</td>
<td>Type C.1.GS</td>
<td>EMWEY_00012830</td>
<td>1948</td>
<td>768</td>
<td>225</td>
<td>1-25</td>
<td>235 (A)</td>
</tr>
<tr>
<td></td>
<td>Type C.1.M6</td>
<td>EMWEY_00012830</td>
<td>1759</td>
<td>768</td>
<td>225</td>
<td>1-25</td>
<td>235 (A)</td>
</tr>
<tr>
<td></td>
<td>Type D.1.GS</td>
<td>EMWEY_00012800</td>
<td>1395</td>
<td>777</td>
<td>258</td>
<td>1-22</td>
<td>233 (G)</td>
</tr>
<tr>
<td></td>
<td>Type D.1.M6</td>
<td>EMWEY_00012800</td>
<td>1170</td>
<td>774</td>
<td>257</td>
<td>1-23</td>
<td>234 (A)</td>
</tr>
<tr>
<td></td>
<td>Type G.1.GS</td>
<td>EMWEY_00012780</td>
<td>1381</td>
<td>771</td>
<td>256</td>
<td>1-23</td>
<td>234 (A)</td>
</tr>
<tr>
<td></td>
<td>Type G.1.M6</td>
<td>EMWEY_00012780</td>
<td>1180</td>
<td>778</td>
<td>248</td>
<td>1-23</td>
<td>234 (A)</td>
</tr>
<tr>
<td></td>
<td>Type J.1.GS</td>
<td>EMWEY_00012770</td>
<td>1176</td>
<td>780</td>
<td>259</td>
<td>1-26</td>
<td>237 (A)</td>
</tr>
<tr>
<td></td>
<td>Type J.1.M6</td>
<td>EMWEY_00012770</td>
<td>1423</td>
<td>882</td>
<td>293</td>
<td>1-26</td>
<td>268 (A)</td>
</tr>
<tr>
<td>Multicopy gene family 4 (Type I SAGs)</td>
<td>Type I.1.GS</td>
<td>EMWEY_00004750</td>
<td>1271</td>
<td>1002</td>
<td>333</td>
<td>1-15</td>
<td>311 (G)</td>
</tr>
<tr>
<td></td>
<td>Type I.3.GS</td>
<td>EMWEY_00004750</td>
<td>1308</td>
<td>978</td>
<td>325</td>
<td>1-15</td>
<td>303 (G)</td>
</tr>
<tr>
<td></td>
<td>Type I.2.M6</td>
<td>EMWEY_00004780</td>
<td>2004</td>
<td>1011</td>
<td>336</td>
<td>1-15</td>
<td>314 (G)</td>
</tr>
<tr>
<td>Single copy gene (Type E SAGs)</td>
<td>Type E.1.GS</td>
<td>EMWEY_00008040</td>
<td>1166</td>
<td>708</td>
<td>235</td>
<td>1-21</td>
<td>213 (V)</td>
</tr>
<tr>
<td></td>
<td>Type E.1.M6</td>
<td>EMWEY_00008040</td>
<td>1166</td>
<td>708</td>
<td>235</td>
<td>1-21</td>
<td>213 (V)</td>
</tr>
<tr>
<td>Single copy gene (Type h SAGs)</td>
<td>Type H.1.GS</td>
<td>EMWEY_00058130</td>
<td>1343</td>
<td>1014</td>
<td>337</td>
<td>1-15</td>
<td>315 (S)</td>
</tr>
<tr>
<td></td>
<td>Type H.1.M6</td>
<td>EMWEY_00058130</td>
<td>1396</td>
<td>1014</td>
<td>337</td>
<td>1-15</td>
<td>315 (S)</td>
</tr>
</tbody>
</table>
The transcripts possessing complete coding sequence (CDS) that produced such translated products were assigned to 1 of 21 surface antigen (SAg) genes following the naming convention of Reid et al. (2013, personal communication) for the nuclear genome assembly and annotation of the *Eimeria maxima* Weybridge (Wey) strain (see Table 5.1). By mapping these 21 SAGs against genome sequence assemblies from *E. maxima* Houghton (H) strain and *E. maxima* Weybridge (Wey) strain available freely online (ftp://ftp.sanger.ac.uk/pub/pathogens/ar11/assemblies/), 32 SAg transcripts of *E. maxima* GS and *E. maxima* M6 were shown to be members of 4 multicopy surface antigen gene families and the remaining 4 SAg transcripts could be easily shown to be similar to single copy genes on the reference genomes (2 loci with 2 examples each). A summary of these 21 SAg transcripts obtained from *E. maxima* GS and *E. maxima* M6 and their products is provided in Table 5.1.

The first, and most complicated, multicopy gene family (Type A SAGs) includes transcripts identified from *E. maxima* GS (6 SAg transcripts) and *E. maxima* M6 (7 SAg transcripts) and representing 8 distinct SAg genes expressed by one or both *E. maxima* strains (Table 5.1). Five genes were expressed by both GS and M6; 1 gene was expressed by GS only; 2 genes were expressed by M6 only (see Figure 5.2). These SAGs were aligned with concatenated genomic contigs of *E. maxima* Houghton (contig06282, contig22518, contig22517, and contig14670). The concatenation of the genomic contigs of *E. maxima* Houghton and Weybridge strains was only possible after 13 SAGs from GS and M6 were aligned to *E. maxima* genome sequence scaffolds (Figure 5.2). The concatenation of genomic sequences of H and Wey strains was necessary because these genome scaffolds are comparatively short sequences; considering the coverage density that was obtained in the present study and the difficulty of forming reliable contigs for the Type A SAg transcripts, it is not surprising that these genome sequencing efforts failed to generate contiguous sequence through this region of the genome. The SAg transcripts and *E. maxima* H genomic sequences could be used to assemble a number of genomic sequence scaffolds being generated as part of a genome sequencing project targeting the *E. maxima* Weybridge strain. The alignment of the coding regions of *E. maxima* GS and *E. maxima* M6 Type A SAg transcripts against homologues of *E. maxima* H and Wey revealed that there is a high level of similarity in the coding regions among these four strains with
limited polymorphisms between SAg orthologs (3 to 35 SNPs per gene; see Table 5.2). However, perhaps not surprisingly given their surface presentation, 6 of 8 Type A SAgs had relatively high $\delta_N/\delta_S$ ratios (1.44 to 2.40) indicating selective pressure on these loci. The percentage identity in the coding regions among the three aligned strains at particular genetic loci was > 90% in most of the aligned sequences (see Table 5.2). To validate transcripts generated from GS and M6, these transcripts were mapped and aligned with genomic sequences of H and Wey strains. The GS and M6 transcripts were mapped to genomic sequences of H and Wey successfully and all genes identified in this large family of multicopy genes had identical structure with a 5'-UTR followed by 4 exons (2 short, 1 long, 1 short) with 3 intervening introns and finally a 3' - UTR (Figure 5.1a). The aligned genomic and transcriptome sequences confirmed that gene structure was conventional with all introns possessing highly conserved motifs starting with nucleotides ‘GT’ at the 5'-end of each intron and ending with a conserved ‘AG’ sequence at the 3'-end of each intron. Some introns possessed an obvious polypyrimidine tract upstream of the latter AG sequence but not all. There were highly conserved regions in the 5'-UTR in all transcripts obtained from *E. maxima* GS and M6 that belonged to this first multicopy gene family, as well as within comparable regions of the genomic sequences from *E. maxima* H and Wey. Likewise, the region of the genes encoding the 3’-end of the 4th and final exon from each SAg gene was highly conserved, particularly in the region of the stop codon for each CDS; see Figure 5.3 for an alignment of all Type A SAg transcripts obtained from GS and M6 as well as orthologs inferred from the genomic sequences of *E. maxima* H and Wey strains.

In contrast to the wide variation observed in the first multicopy gene family, the next multicopy family of surface antigens demonstrated high levels of sequence identity at each genetic locus with all four strains of *E. maxima* for which sequence data were available. Multicopy gene family 2 (Type B SAgs) consists of a total of 8 SAg transcripts identified in *E. maxima* GS (4 transcripts) and *E. maxima* M6 (4 transcripts). These 8 SAg transcripts have homology with 4 SAg genes of *E. maxima* Wey and H strains (e.g. genes EMWEY_00005470, EMWEY_00005480, EMWEY_00005500 and EMWEY_00005510). Our 8 SAgs aligned with genomic sequence of *E. maxima* H strain (contig00278) and *E. maxima* Wey (scaffold30) with high similarity (0-2 non-synonymous changes in aa) in the coding region, 5'UTR, and 3'UTR. Within this multicopy gene family, there was little to
no strain variation (99.8-99.9% sequence identity) among coding regions from a single locus from four different *E. maxima* strains (M6, GS, H and Wey). In contrast, there was considerable sequence diversity among loci; for example, there was only 72.6% aa identity between the propeptides generated from loci EMWEY_00005470 and EMWEY_00005480 (Figure 5.4). Despite exhaustive attempts to assemble Illumina reads against reference Type B.3 CDS sequences, Type B.3 was expressed by neither *E. maxima* GS nor *E. maxima* M6 but was present in both reference genome sequences. Type B.3 gene might be stage specific gene that is expressed during merogony or gametogony but this gene is not expressed during sporogony.

Multicopy gene family 3 (Types C, D, G, and J) consists of 4 SAg genes that were each transcribed by both *E. maxima* GS and *E. maxima* M6 (Table 5.1). The alignment of these 8 SAg transcripts with genomic sequence of *E. maxima* Wey (scaffold 93) identified these transcripts as belonging to following genes: EMWEY_00012770; EMWEY_00012780; EMWEY_00012800; and EMWEY_00012830 (see Table 5.1). These 8 SAg transcripts from *E. maxima* GS and M6 could also be aligned with genomic sequences of *E. maxima* H (contig01967, contig02104, and contig09197). The percentage of identical sites in the coding regions among of *E. maxima* GS, *E. maxima* M6, and *E. maxima* H, and *E. maxima* Wey strains was (87.8% - 100%) among all the aligned SAg (Figure 5.5).
Figure 5.2 Multicopy Gene Family 1 - Type A Surface Antigens. There were 13 GPI anchored proteins identified in *E. maxima* GS and *E. maxima* M6. These SAgs were aligned with genomic sequences of *E. maxima* Weybridge strain (many scaffolds, blue); and *E. maxima* Houghton strain (many contigs - blue). The coding regions of *E. maxima* GS SAgs (blue) and *E. maxima* M6 SAgs (red) are mapped below reference genomic sequences from *E. maxima* H and Wey; coding regions (yellow), UTRs (grey) and genes (green) are annotated on the genomic reference sequences. Type A.2.G was expressed only in GS strain. Type A.5.M6 and Type A.7.M6 were expressed only in M6 strain.
Figure 5.3 Alignment of Type A SAg transcripts of *Eimeria maxima* GS and M6 aligned with mRNA transcripts inferred from genomic sequences of *E. maxima* H and Wey strains. The coding regions of *E. maxima* GS SAGs (blue), *E. maxima* M6 SAGs (red), *E. maxima* H SAGs (yellow) and *E. maxima* Wey SAGs (green) are indicated as well as 5' and 3' UTRs (grey). Although there was considerable diversity among the various transcripts (concentrated primarily in exon 3), there were highly conserved regions near the end of the 5'-UTR and the start of the first exon, in part of the third exon (~540-700, above) and in the region of the stop codon for each CDS.
Figure 5.4 Multi-copy gene family Type B. This family consists of 5 SAg genes identified in *E. maxima* M6 and GS strains; 4 of the SAg genes were each expressed by the *E. maxima* M6 and GS strains but Type B.3 was not expressed by either strain during sporogony. These SAggs were aligned with genomic sequences of *E. maxima* Weybridge strain scaffold30 (blue); and *E. maxima* Houghton strain (contig00278-blue). SAg mRNA sequences (burgundy) and the coding regions of *E. maxima* GS SAggs (blue) and *E. maxima* M6 SAggs (red) are mapped below reference genomic sequences from *E. maxima* H and Wey; coding regions (yellow), UTRs (grey) and genes (green) are annotated on the genomic reference sequences.
**Figure 5.5** Multi-copy gene family Type C, D, J, and G. This family consists of 4 SAg genes, each expressed by the *E. maxima* M6 and GS strains. These SAGs were aligned with genomic sequences of *E. maxima* Weybridge strain scaff93 (blue); and *E. maxima* Houghton strain (contigs 01967, 02104, and 09197 - blue). SAG mRNA sequences (burgundy) and the coding regions of *E. maxima* GS SAGs (blue) and *E. maxima* M6 SAGs (red) are mapped below reference genomic sequences from *E. maxima* H and Wey; coding regions (yellow), UTRs (grey) and genes (green) are annotated on the genomic reference sequences.

Multicopy gene family 4 (Type I SAGs) include 3 SAG genes expressed in *E. maxima* GS and *E. maxima* M6 (Table 5.1). These three SAG transcripts aligned with genomic sequence of *E. maxima* Wey (scaff1036). Four contigs (contig04120, contig09147, contig14671 and contig14529) of *E. maxima* H genomic sequences aligned against scaff1036 of Wey strain to create a reference for mapping our SAG transcripts. The alignment of 3 SAG transcripts identified in GS and M6 strains against genomic sequences of Wey and H strains discovered high sequence identities (99.4% - 100.0%) in the coding region sequences among the aligned strains (GS, M6, Wey, and H) at each gene locus. However, among this multicopy gene family, 2 SAGs (Types I.1 and I.3) were expressed in GS strain only and 1 SAG was expressed in M6 (Type I.2) only (Figure 5.6).
Figure 5.6 Multi-copy gene family Type I. This family consists of 2 SAg genes (Type I.1.GS and Type I.3.GS) expressed in *E. maxima* GS strains, and 1 SAg gene (type I.2.M6) was expressed in *E. maxima* M6 only. These SAggs were aligned with genomic sequences of *E. maxima* Weybridge strain scaff1036 (blue); and *E. maxima* Houghton strain (contigs 04120, 09147, 14671, and 14529 - blue). SAg mRNA sequences (burgundy) and the coding regions of *E. maxima* GS SAgs (blue) and *E. maxima* M6 SAgs (red) are mapped below reference genomic sequences from *E. maxima* H and Wey; coding regions (yellow), UTRs (grey) and genes (green) are annotated on the genomic reference sequences.

It was also found that four SAgs were not part of a multi-copy gene family. These SAgs were Type E.1.GS, Type E.1.M6, Type H.1.GS, and Type H.1.M6 These SAgs aligned with genomic sequences of *E. maxima* Houghton and *E. maxima* Weybridge strains. Type E and Type H SAgs demonstrated high sequence identity at the nucleotide level in the coding regions (~99.9-100% pairwise sequence identity) among the aligned SAgs of GS, M6, H, and Wey (see Figures 5.7 & 5.8, respectively).
Figure 5.7 Single copy gene (Type E) encoded of 2 SAgs identified in GS and M6 strains. The SAg mRNA sequences (burgundy), UTRs (gray) and coding regions of *E. maxima* GS SAgs (blue) and *E. maxima* M6 SAgs (red), Houghton (yellow), and Weybridge (yellow) were aligned. All four strains of *E. maxima* had identical sequence in the coding regions (100% - see green identity bar at top of figure; only a single base within the 3’UTR demonstrated any variation among the strains). The two SAgs of GS and M6 aligned with scaff49 and contig0304 of Weybridge and Houghton strains respectively; the Type E genes (green) are annotated on the genomic reference sequences.

Figure 5.8 Single copy gene (Type H) consisting of 2 SAgs identified in GS and M6 strains. The SAgs mRNA sequences (burgundy), UTRs (gray) and coding regions of *E. maxima* GS SAgs (blue) and *E. maxima* M6 SAgs (red), Houghton (yellow), and Weybridge (yellow) were aligned. A high degree of identity in coding regions (99.9-100% pairwise sequence identity) was noted among the aligned sequences. The two SAgs of GS and M6 aligned with Weybridge (scaffold2344) and Houghton (contigs 07399 and 09587); the Type H genes (green) are annotated on the genomic reference sequences.
The coding regions identified in each of the 36 transcripts (representing 21 distinct SAg genes) were translated into peptides. All 36 peptides were determined to have putative signal peptides, GPI-anchor cleavage sites, and terminal transmembrane domains. The positions of GPI-anchor cleavage sites were at different amino acids among the identified SAggs. Most of the SAggs showed that GPI cleavage site was at amino acid G; but 2 peptides showed the GPI cleavage site at amino acid V. The prediction of transmembrane helices performed using Hidden Markov Model (TMHMM) discovered a total of 36 peptide sequences containing at least a single transmembrane domain (see Table 5.1).

Primary translation products (propeptides) of the 13 expressed Type A SAg transcripts (7 from M6 and 6 from GS) had generally conserved structures, particularly in the signal peptide region (residues 1 to ~20-23, 92% pairwise identity) and a neighbouring highly conserved region (residues 31 to 57, 97% pairwise identity) in what will become the terminus of the mature peptide after cleavage of the signal peptide and acquisition of the GPI anchor (Figure 5.8) and there was an additional region of high sequence similarity (residues 96 to 164, ~90% pairwise sequence identity over the entire region). By mapping the exons that encoded the various propeptide translation products onto the alignment, it became clear that the 5’-ends of exons 2 and exon 4 were highly conserved (residues 33-55 and residues 340-351) across all Type A SAggs. Several of these regions of conserved sequences were associated with predicted antigenic regions (Figure 5.9).
Figure 5.9 Aligned and annotated Type A surface antigen predicted translation products from *Eimeria maxima* GS and M6. The transcript coding regions of *E. maxima* GS SAgS (blue) and *E. maxima* M6 SAgS (red) are indicated below each translated propeptide. Conserved sequences were found in the signal peptide region (residues 1 to ~20-23) as well as in regions predicted to possess antigenic epitopes.

Analysis of the 36 GPI-anchored SAGs revealed considerable differences in the relative abundance of transcripts from each SAg gene based on their representation in sequences obtained from cDNA libraries prepared from purified sporulating (18 h, 22 h, 38 h) and fully sporulated sporozoites of *E. maxima* GS and *E. maxima* M6. An indirect measure of the relative abundance of various SAg transcripts was determined by assembling all available paired reads against coding regions (CDS’s) of all identified SAg mRNA sequences using Geneious (see materials and methods for details). In general, it was noticed that GPI-anchored SAg transcripts that are transcribed by both strains have more transcripts than SAgS that are expressed in one strain but are not expressed in the
other strain. For instance, the highest transcript abundance was found for Type E.1 SAg transcripts which were transcribed abundantly in both strains (GS and M6). The number of reads assembled to the coding region of the Type E.1 SAg transcripts from *E. maxima* GS and *E. maxima* M6 were 508,635 and 25,762 reads, respectively. In contrast, the SAg transcript with the lowest apparent abundance of transcripts was reported with SAg EMWEY_00047560 which is expressed only in *E. maxima* GS (Figure 5.10).

Unique GPI-anchored SAgs were identified in *E. maxima* GS and *E. maxima* M6 and vice versa. There were 3 unique GPI anchored SAgs (Type A.2.GS, Type I.3.GS, and Type I.1.GS) that were found to be expressed in *E. maxima* GS but not in *E. maxima* M6. Likewise, there were 3 GPI anchored SAgs (Type A.5.M6, Type A.7.M6, and Type I.2.M6) that were expressed in *E. maxima* M6 but were not expressed in *E. maxima* GS. The uniquely expressed SAgs in both strains were part of multi-copy gene families (Type A and Type I multi-copy gene family). The abundance of transcripts of the unique SAgs was less than the abundance of transcripts of SAgs shared between *E. maxima* GS and *E. maxima* M6 (Figure 5.10).
Figure 5.10 the abundance of gene expression of 36 GPI-anchored SAgS identified in *E. maxima* GS (18 SAgS) and *E. maxima* M6 (18 SAgS). The blue columns represent SAgS of GS strain. The red columns represent SAgS of M6 strain. The highest abundance of gene expression was discovered in Type E.1.M6 but the lowest abundance of gene expression was reported in Type G.1.M6. The blue striped columns represent the 3 SAgS expressed in GS strain only. The red striped columns (3 SAgS) showed SAgS expressed M6 strain only. It was clear that many of the SAg transcripts that are shared between GS and M6 have higher abundance of expression than SAgS that were expressed in one strain only.
Table 5.2 Comparison of coding regions of 36 GPI-anchored SAgs from *Eimeria maxima* GS and M6 strains with Houghton, and Weybridge strains. Among the compared SAgs, the coding regions were identical to nearly identical at most, but not all, loci.

<table>
<thead>
<tr>
<th>Transcript (SAg)</th>
<th>Shared or unique</th>
<th>Relative Expression of CDS* (=RPKM/10)</th>
<th>Pairwise Sequence Identity (vs H)</th>
<th>Pairwise Sequence Identity (vs Wey)</th>
<th>Pairwise Sequence Identity (GS vs M6)</th>
<th>Polymorphism of CDS’s (δN/δS ratio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type A.1 GS</td>
<td>Shared</td>
<td>1.48</td>
<td>100.0%</td>
<td>99.5%</td>
<td>97.2%</td>
<td>6 SNPs</td>
</tr>
<tr>
<td>Type A.1 M6</td>
<td>Shared</td>
<td>2.04</td>
<td>97.2%</td>
<td>97.3%</td>
<td>N/A</td>
<td>12 SNPs</td>
</tr>
<tr>
<td>Type A.2 GS</td>
<td>Unique</td>
<td>5.15</td>
<td>99.8%</td>
<td>98.6%</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Type A.3 GS</td>
<td>Shared</td>
<td>4.96</td>
<td>99.5%</td>
<td>99.4%</td>
<td>98.8%</td>
<td>8 SNPs</td>
</tr>
<tr>
<td>Type A.3 M6</td>
<td>Shared</td>
<td>2.32</td>
<td>98.9%</td>
<td>99.4%</td>
<td>98.6%</td>
<td>17 SNPs</td>
</tr>
<tr>
<td>Type A.4 GS</td>
<td>Shared</td>
<td>2.54</td>
<td>99.5%</td>
<td>97.0%</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Type A.4 M6</td>
<td>Shared</td>
<td>1.62</td>
<td>98.6%</td>
<td>97.8%</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Type A.5 M6</td>
<td>Unique</td>
<td>1.90</td>
<td>98.4%</td>
<td>99.4%</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Type A.6 GS</td>
<td>Shared</td>
<td>8.81</td>
<td>98.1%</td>
<td>95.7%</td>
<td>98.9%</td>
<td>35 SNPs</td>
</tr>
<tr>
<td>Type A.6 M6</td>
<td>Shared</td>
<td>5.31</td>
<td>98.9%</td>
<td>95.8%</td>
<td>N/A</td>
<td>3 SNPs</td>
</tr>
<tr>
<td>Type A.7 M6</td>
<td>Unique</td>
<td>0.59</td>
<td>99.6%</td>
<td>99.8%</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Type A.8 GS</td>
<td>Shared</td>
<td>5.79</td>
<td>99.0%</td>
<td>98.1%</td>
<td>98.7%</td>
<td>21 SNPs</td>
</tr>
<tr>
<td>Type A.8 M6</td>
<td>Shared</td>
<td>2.64</td>
<td>99.0%</td>
<td>98.1%</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Type B.1 GS</td>
<td>Shared</td>
<td>11.85</td>
<td>99.9%</td>
<td>99.9%</td>
<td>100.0%</td>
<td>1 SNPs</td>
</tr>
<tr>
<td>Type B.1 M6</td>
<td>Shared</td>
<td>10.99</td>
<td>99.9%</td>
<td>99.9%</td>
<td>100.0%</td>
<td>2 SNPs</td>
</tr>
<tr>
<td>Type B.2 GS</td>
<td>Shared</td>
<td>0.40</td>
<td>99.9%</td>
<td>99.7%</td>
<td>100.0%</td>
<td>1 SNPs</td>
</tr>
<tr>
<td>Type B.2 M6</td>
<td>Shared</td>
<td>0.80</td>
<td>99.9%</td>
<td>99.7%</td>
<td>100.0%</td>
<td>2 SNPs</td>
</tr>
<tr>
<td>Type B.3 GS</td>
<td>Shared</td>
<td>17.38</td>
<td>99.9%</td>
<td>99.9%</td>
<td>100.0%</td>
<td>1 SNPs</td>
</tr>
<tr>
<td>Type B.4 M6</td>
<td>Shared</td>
<td>12.36</td>
<td>99.9%</td>
<td>99.9%</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Type B.5 GS</td>
<td>Shared</td>
<td>11.96</td>
<td>99.8%</td>
<td>99.8%</td>
<td>99.9%</td>
<td>2 SNPs</td>
</tr>
<tr>
<td>Type B.5 M6</td>
<td>Shared</td>
<td>1.44</td>
<td>99.9%</td>
<td>99.9%</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Type C.1 GS</td>
<td>Shared</td>
<td>50.68</td>
<td>98.7%</td>
<td>98.7%</td>
<td>98.7%</td>
<td>10 SNPs</td>
</tr>
<tr>
<td>Type C.1 M6</td>
<td>Shared</td>
<td>21.06</td>
<td>98.7%</td>
<td>100.0%</td>
<td>98.7%</td>
<td>10 SNPs</td>
</tr>
<tr>
<td>Type D.1 GS</td>
<td>Shared</td>
<td>5.75</td>
<td>100.0%</td>
<td>87.8%</td>
<td>87.8%</td>
<td>87 SNPs</td>
</tr>
<tr>
<td>Type D.1 M6</td>
<td>Shared</td>
<td>0.12</td>
<td>87.8%</td>
<td>100.0%</td>
<td>87.8%</td>
<td>87 SNPs</td>
</tr>
<tr>
<td>Type G.1 GS</td>
<td>Shared</td>
<td>0.40</td>
<td>100.0%</td>
<td>72.4%</td>
<td>72.6%</td>
<td>138 SNPs</td>
</tr>
<tr>
<td>Type G.1 M6</td>
<td>Shared</td>
<td>0.07</td>
<td>72.6%</td>
<td>99.6%</td>
<td>72.6%</td>
<td>138 SNPs</td>
</tr>
<tr>
<td>Type J.1 GS</td>
<td>Shared</td>
<td>13.53</td>
<td>100.0%</td>
<td>97.1%</td>
<td>85.8%</td>
<td>22 SNPs</td>
</tr>
<tr>
<td>Type J.1 M6</td>
<td>Shared</td>
<td>0.71</td>
<td>85.8%</td>
<td>100.0%</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Type I.1 GS</td>
<td>Unique</td>
<td>0.62</td>
<td>99.4%</td>
<td>100.0%</td>
<td>6 SNPs</td>
<td></td>
</tr>
<tr>
<td>Type I.3 GS</td>
<td>Unique</td>
<td>2.45</td>
<td>99.2%</td>
<td>99.2%</td>
<td>N/A</td>
<td>2 SNPs</td>
</tr>
</tbody>
</table>

*Note:* Relative Expression of CDS refers to the expression level relative to RPKM/10.
<table>
<thead>
<tr>
<th>Transcript (SAg)</th>
<th>Shared or unique</th>
<th>Relative Expression of CDS(^a) (=RPKM/10)</th>
<th>Pairwise Sequence Identity (vs H)</th>
<th>Pairwise Sequence Identity (vs Wey)</th>
<th>Pairwise Sequence Identity (GS vs M6)</th>
<th>Polymorphism of CDS’s (δ(N_2)/δ_3) ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I.2.M6</td>
<td>Unique</td>
<td>0.28</td>
<td>100.0%</td>
<td>100.0%</td>
<td>N/A</td>
<td>1 SNPs (δ_0/δ_3=0.00) (HWM)</td>
</tr>
<tr>
<td>Type E.1.GS</td>
<td>Shared</td>
<td>34.99</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
<td>0 SNPs (δ_0/δ_3=n/a) (HWGM)</td>
</tr>
<tr>
<td>Type E.1.M6</td>
<td>Shared</td>
<td>191.07</td>
<td>100.0%</td>
<td>100.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type H.1.GS</td>
<td>Shared</td>
<td>0.67</td>
<td>99.9%</td>
<td>N/A partial</td>
<td>100.0%</td>
<td>0 SNPs (δ_0/δ_3=n/a) (HWGM)</td>
</tr>
<tr>
<td>Type H.1.M6</td>
<td>Shared</td>
<td>2.05</td>
<td>99.9%</td>
<td>N/A partial</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 5.5 Discussion

The two closely related *E. maxima* strains (GS and M6) examined in this study are known to lack immunological cross-protection but are otherwise highly similar biologically (Martin et al. 1997; Barta et al. 1998; Al-Badri and Barta 2012), and genetically (Barta et al. 1998; Ogedengbe et al. 2011). Despite these similarities, there is a marked lack of immunological cross-protection when chickens are infected with one strain and then challenged 2 weeks later with the other (Martin et al 1997; Beattie et al. 2001); birds similarly infected with one of the two strains demonstrate complete, sterile immunity to reinfection by the homologous strain when challenged similarly (Beattie 1997; Beattie et al. 2001). One of the key findings of the studies by Beattie (1997) and Beattie et al. (2001) was that immunity against challenge with either strain of *E. maxima* was directed functionally against the invading sporozoites. Whatever antigenic differences that exist between these two strains and contribute to the lack of immunological cross-protection must be expressed at the level of the invading sporozoite, although perhaps not exclusively so. Therefore, the approach of the present study was to compare transcribed genes between these two strains with an emphasis on structural or antigenic molecules such as surface antigens that might differ between these two strains. Families of surface expressed, GPI-anchored antigens are found commonly on the surface of apicomplexan zoites such as *Eimeria tenella* (see Taberés et al. 2004), *Toxoplasma gondii* (see Jung et al. 2004), *Sarcocystis neurona* (see Howe et al. 2005) and *Plasmodium falciparum* (see Gilson et al. 2006). Attachment and invasion of host cells by apicomplexan parasites (such as *T. gondii*) is dependent on SAGs (see Lekutis et al. 2001) and other molecules so recognition of critical SAGs immunologically may permit control of an infection by the host.
The examination of the transcriptome of two closely related strains of *E. maxima* demonstrated at least 36 GPI-anchored SAg transcripts in *E. maxima* GS (18 SAg transcripts) and *E. maxima* M6 (18 SAg transcripts). The 36 identified transcripts encoding putative SAsgs were associated with at least 21 distinct SAg genes with apparent homology with SAg genes of *E. maxima* Wey and H. The genome of *E. maxima* Wey has recently been sequenced and annotation of its genes largely completed so the gene nomenclature of Reid et al. (2013, personal communication) has been followed when referring to these SAg genetic loci. Most (32) of the SAg transcripts were members of 4 multi-gene families and the 4 remaining SAg transcripts were associated apparently with isolated (single) genes.

The identified transcripts encoded putative SAg propeptides that all have putative signal peptides, GPI-cleavage sites, and transmembrane structures typical of GPI-anchored surface expressed proteins (Tabarés et al. 2004). The comparison of aligned coding regions among the identified SAsgs showed tremendous diversity in the sequences in both the mRNA (i.e. transcript contigs) and amino acid translation products among *E. maxima* GS, *E. maxima* M6, *E. maxima* Wey and *E. maxima* H. Surprisingly, most expressed SAg loci were highly similar among all strains (up to 4) for which sequence data were available, including members of the large multicopy gene family that encodes Type A SAg transcripts. However, there were differentially expressed or unique GPI-anchored SAsgs that were expressed solely in *E. maxima* GS (3 SAg transcripts) but were not expressed by *E. maxima* M6. Likewise, 3 GPI-anchored SAsgs were found to be expressed only by *E. maxima* M6. These observations suggest that there is a distinctive repertoire of GPI-anchored SAg transcripts expressed by each of the two *E. maxima* strains (GS and M6).

The abundance of expression of the 36 SAg transcripts varied greatly among particular genes and between strains. Although the abundance of SAg transcripts was not measured directly, the relative abundance of Illumina paired-reads that mapped to the coding region of a particular SAg transcript sequence using highly sensitive assembly parameters was used as an indirect measure of transcript abundance in the original mRNA samples prepared from each strain. Interestingly, the highest level of expression was discovered between SAg transcripts that were expressed by both *E. maxima* strains.
whereas the lowest abundance of expression was discovered in SAg transcripts that were unique to one strain. This suggests that these ‘minor’ antigens may be more important antigenic targets for the generation of a protective immune response than highly expressed SAg (potentially, ‘immunodominant’ antigens) that would seem natural targets of an immune response by the host. This may also explain why cloning of immunodominant antigens (usually through selection of cDNA clones using hyperimmune IgG or IgY from chickens that had recovered from an *Eimeria* sp. infection) has not produced subunit vaccines that effectively protect against even homologous coccidial challenge infections (e.g. Jenkins and Dame 1987; Kopko et al. 2000; Danforth et al. 1989; Jenkins et al. 1991; Shirley et al. 2007; Chapman et al. 2013).

The differential expression of Type I SAg genes by *E. maxima* GS and M6 (see Figure 5.5, Table 5.1) was particularly interesting. Types I.1 and I.3 were expressed by GS only and Type I.2 was expressed by M6 exclusively. However, comparison of these transcripts with the genomic sequences from *E. maxima* Wey and H strains indicate that all three genes (encoding SAg Type I.1, Type I.2 and Type I.3) are selectively expressed in GS and M6 with high pairwise identity at both the nucleotide and amino acid level (98.8% identical - 4 variable positions, 100% identical - 0 variable positions, and 99.7% identical - 1 variable position, respectively). The presence of all three genes in both reference genomes suggests that the strain specific expression of the Type I SAgS may be selective expression rather than genome deletion of the associated genes. Such selective expression of well-conserved genetic loci may represent one means of expressing a strain-specific antigenic repertoire to the host. Further research is warranted on the Type I SAgS to determine if this is an important genetic locus for antigenic differentiation among these strains.

The interaction of glycosylphosphatidylinositol (GPI)-anchored surface antigens (SAgs) with their avian host has been investigated in a number of *Eimeria* species. Chow et al. (2010) expressed 10 of the 19 known SAgS (Taberés et al. 2004) of *Eimeria tenella* as recombinant fusion proteins in *Escherichia coli*. The N-terminal hydrophobic signal and C-terminal hydrophobic GPI-cleavage regions were removed from each SAg sequence prior to plasmid cloning and bacterial expression so that only the mature peptides, as found
on the surface of the parasite, were expressed as a fusion protein. Chow et al. (2010) found that high levels of macrophage nitric oxide production and IL-1b mRNA transcription were stimulated after treating macrophages with recombinant SAgs 4, 5 or 12. Jahn et al. (2009) identified a surface antigen in *E. tenella* using 2D gel electrophoresis followed by mass spectrometry (EtSAG1). The invasion process of sporozoites was reduced in the presence of monoclonal antibody (2H10E3) raised against EtSAG1. These and other studies (e.g. Jung et al. 2004; Taberés et al. 2004; Howe et al. 2005; Gilson et al. 2006) suggest that there is a vital role for GPI-anchored SAgs in the invasion of host cells and in the induction of immune responses including inflammatory reactions (e.g. Chow et al. 2010) and antibody production (e.g. Howe et al. 2005).

Twenty three GPI-anchored proteins were identified in *Eimeria tenella* when sporozoite and second generation merozoite stages were examined by Tabarés et al. (2004). These SAgs were classified into two multigene families, A and B, based on the positions of six conserved cysteines in the protein alignment. The number of the SAgs that were identified in *E. maxima* GS (18 SAgs), and *E. maxima* M6 (18 SAgs) was similar to the number found in *E. tenella* and these belonged to 4 multicopy gene families. It is well known that *E. maxima* is more variable immunologically than other *Eimeria* species (e.g. Fitz-Coy 1992; Barta et al. 1998; Lew et al. 2003; Cantacessi et al. 2008; Schwarz et al. 2009) and this may be explained partially by the large number of multi-copy gene families encoding SAgs detected in *E. maxima* in the present study and in genome sequencing (Reid et al. 2013, in prep).

The proteomes of different stages of the life cycle of *E. tenella* (unsporulated oocysts, sporulated oocysts, sporozoites, and second-generation merozoite) were compared by Lal et al. (2009). There were 68 GPI-anchored surface proteins identified in the merozoites (47 of which were identified as SAgs) but only 16 GPI-anchored surface proteins (only 4 of which were identified as SAgs) were identified in the sporozoites of *E. tenella*. However, the authors note that the majority of these 16 GPI-anchored, secretory proteins may be SAgs even though they were only able to positively identify 4. A total of 21 putative SAg genes were expressed during sporogony of *E. maxima* GS (expressing 18 SAgs) and *E. maxima* M6 (expressing 18 SAgs). If the GPI-anchored surface expressed
proteins of *E. tenella* detected by Lal et al. (2009) were all SAGs, it would suggest that about 16-20 SAGs are expressed by sporozoites of avian *Eimeria* species. The observations in the present and other studies suggest that more numerous SAGs may be produced during merogonic development than during sporogony and production of sporozoites of *Eimeria* species.

Finally, the observations in the present study support the conclusion that there are strain-specific differences in the repertoire of GPI-anchored surface antigen expression. Highly expressed SAg transcripts (determined through pair-read abundance in the shotgun sequencing of cDNA from both strains generated from poly-adenylated mRNA) were always encoded by genetic loci that were expressed in both *E. maxima* GS and M6. In contrast, GPI-anchored SAg transcripts expressed uniquely by one of the two *E. maxima* strains were expressed at a comparatively modest level compared to shared SAg transcripts. This suggests that strain specific immunologic recognition may be dependent on “minor” SAGs rather than abundant SAGs. The documented differences among the GPI-anchored SAGs expressed by the GS and M6 strains may explain the lack of immunological cross-protection that has been demonstrated between these parasite strains. Using the transcriptome data obtained in the present study to identify and quantify transcription of putative SAGs in these two strains may provide the necessary information to rationally approach development of a subunit vaccine against coccidiosis that could be less costly than live vaccines that are now available for poultry.
Chapter 6 – General Discussion

*Eimeria maxima* GS and *E. maxima* M6 have been the subjects of several studies because these two strains have shown to be immunologically distinct and an infection with one strain does not provide protective immunity against the other strain (Beattie et al. 2001; Martin et al. 1997) even in birds solidly immune to homologous challenge. This is despite the fact that these parasites belong to same *Eimeria* species; both strains have been confirmed to be *E. maxima* by a variety of molecular methods (see Barta et al. 1998; Ogedengbe et al. 2011).

Sporozoites are known to be the primary target of a protective immune response in both of these *E. maxima* strains (Beattie 1997; Beattie et al. 2001) but there has been ambiguity about the molecules possessed or expressed by sporozoites that stimulate the avian immune system. The importance of sporozoites as a stage of *Eimeria* species that can elicit a protective immune response is not limited to *E. maxima*. Jeffers and Long (1985) found that sporozoites of *E. tenella* inhibited with the anticoccidial drug decoquinate stimulated significant protective immunity in chickens subsequently challenged with decoquinate resistant *E. tenella* sporozoites. The elicited immunity eliminated the decrease in weight gain but provided only slight protection against intestinal lesions when treated birds were challenged. Similarly, using decreasing doses of gamma-irradiation to render sporozoites dead, inhibited after cell penetration or inhibited after some intracellular development, Jenkins et al. (1991a) found that protective immune responses were elicited only after invasion by *E. tenella* sporozoites. Dead sporozoites were unable to elicit a protective immune response.

The objectives of this study were to characterize differences between the transcriptomes of the sporogonic stages of *E. maxima* GS and *E. maxima* M6 with the hope to identify variation in the expression of surface antigens (SAgs) between these strains that might explain the lack of cross-immunity between these two parasites in chickens. As part of this comprehensive study of the differences and similarities between these two immunologically distinct strains of *E. maxima* GS and M6, the pattern of oocysts shedding and kinetics of sporulation were studied to understand the morphological changes during sporulation and then to determine the best time for collecting RNA samples. Sporulation could be described to pass through at least 5 morphologically distinguishable stages that
predominated the sporulation process at the following times: unsporulated oocysts at 0 h; sporoblast anlagen at 18 h; sporoblasts without sporocyst walls at 22 h; sporocysts without mature sporozoites at 38 h and finally fully mature oocysts that first appear at 48 h and all oocysts were sporulated by 60 h of sporulation (Al-Badri and Barta 2012). Also, both strains had a prepatent period of approximately 120 h followed by peak oocyst shedding at 144-150 h post inoculation. Mean total oocyst output determined for each strain demonstrated that the fecundity of the M6 strain $(12.8 \times 10^3 \pm 1.95)$ of *E. maxima* was roughly twice that of the GS strain $(6.9 \times 10^3 \pm 3.33)$ when birds were inoculated same dose $(2 \times 10^3$ oocysts/bird).

The result of the second study was that the functional annotation of 10,000 transcripts generated from *de novo* assembly discovered putative surface antigen in *E. maxima* GS (19 SAgs) and *E. maxima* M6 (18 SAgs). It was also found that most of transcripts (~60%) have metabolic and cellular process function. Our transcripts also showed that they have similar matches with other protozoan parasites. Out of 5,000 contigs of GS strain, 2,898 transcripts showed that they have similar matches with other species of protozoa (especially *T. gondii*); and 2,569 contigs out of 5,000 transcripts of M6 strain showed that they have similar matches with other protozoan parasites (again, principally *T. gondii*).

A large number of transcripts (36) was identified in the third study that likely encode GPI-anchored surface proteins (likely surface antigens, SAgs) in *E. maxima* GS (18 SAgs) and *E. maxima* M6 (18 SAgs). The biosynthesis of GPI anchored proteins is known to occur in the endoplasmic reticulum (ER). The precursor proteins go through post-translation modifications that lead to the addition of the GPI to the carboxyl-termini of the proteins. The newly synthesized GPI-anchored proteins are transported to the surface of parasites where they can function as antigens, adhesive structures or for other purposes (e.g. Ferguson 1988; Ikezawa 2002; see Figure 6.1).
The 36 identified SAg transcripts demonstrated sequence similarities with 21 named SAg genes of *E. maxima* Weybridge strain (Reid et al. 2013, personal communication). These 36 SAGs were grouped into 4 multi-gene families and 2 single-gene families after they were aligned against *E. maxima* Houghton strain genomic contigs that are available publicly (http://www.genomemalaysia.gov.my/emaxdb/). The translations of identified SAGs showed that they all have signal peptides, GPI cleavage sites, and transmembrane structures (see Figure 5.1).

Pairwise comparisons among the coding regions of the identified SAGs showed varying degrees of diversity in both the DNA and resultant hypothetical amino acids translation products among *E. maxima* strains (GS, M6, H, and Wey). Analysis of SNPs from SAg orthologs from GS, M6, Wey, and H strains demonstrated a wide range of $\delta_N/\delta_S$ ratios (see Table 5.2); in general, the highly expressed genes were comparatively conserved with $\delta_N/\delta_S$ ratios of $\leq 1.0$, indicating low selective pressure (Yang and Bielawski 2000). The top 5 most expressed SAg loci (Types E, C, B.4, B.1 and J) all had $\delta_N/\delta_S$ ratios of $\leq 0.69$; indeed, the 4 most highly expressed loci had $\delta_N/\delta_S$ ratios of 0 despite having 12 SNPs among them. In contrast, some SAg transcripts expressed at low levels demonstrated...
relatively higher selective pressure ($\delta_N/\delta_S$ ratios was »1.0). In the case of the least expressed SAg transcript, Type G.1, the $\delta_N/\delta_S$ ratio was a comparatively high 2.37. Type A SAg loci demonstrated $\delta_N/\delta_S$ ratios that suggest many members of this particular multicopy gene family are under positive selective pressure (Table 5.2). The observed polymorphisms between some GS and M6 shared SAgS and high $\delta_N/\delta_S$ ratios for some of these transcripts suggest that these genes may participate in the biologically expressed lack of cross-immunity between these two parasite strains.

In addition to polymorphisms within particular loci, differential expression of SAg transcripts may also be contributing to the lack of immunological cross-reactivity of these two *E. maxima* strains. Three SAgS expressed in *E. maxima* GS were not expressed in *E. maxima* M6; and 3 SAgS were expressed only in *E. maxima* M6. Similar to transcripts with comparatively higher $\delta_N/\delta_S$ ratios, the highest levels of transcript expression were observed for SAgS that were shared (expressed) by both the M6 and GS strains whereas the lowest abundance of expression was frequently observed in SAgS that were unique to one strain or the other. This suggests that these ‘minor’ antigens may be more important antigenic targets for the generation of a protective immune response than highly expressed (potentially ‘immunodominant’ antigens) SAgS that would seem natural targets of an immune response by the host.

In the present study, no direct evidence was obtained to confirm that any of these SAg transcripts were successfully surface expressed nor was evidence obtained to confirm the antigenicity of these potential SAg molecules and to determine any strain-specificity of such an immune response. Testing the immunogenicity of particular SAgS is complicated with *Eimeria* spp. because a protective immune response is largely T-cell mediated and therefore comparatively simple serum antibody ELISAs are unlikely to be informative. The difficulty of generating knockout parasites because of the lack of an *in vitro* culture system (Chapman et al. 2013) and complications of subsequently evaluating any impacts on cross-protection of such genetically manipulated lines of parasites *in vivo* hamper attempts to directly measure the role of particular SAgS in poultry *Eimeria*.

Glycosylphosphatidylinositol (GPI)-anchored proteins (some of which may be antigens) are expressed on the surface of apicomplexan zoites stages of life cycle and many other eukaryotic cells. These surface antigens (SAgs) are well known to elicit host immune
responses and also have important roles in apicomplexan parasite biology affecting pathogenicity such as host cell invasion and immune evasion or activation (Lekutis et al. 2001; Jung et al. 2004; Tabarés et al. 2004; Gilson et al. 2006; Wasmuth et al. 2012). Diverse sets of GPI-anchored proteins (with apparent homology to those of T. gondii) have been identified in other members of the Apicomplexa such as E. tenella (see Tabarés et al. 2004). Several studies implicate apicomplexan GPI-anchored proteins in immune system activation or stimulation. For instance, Schofield and Hackett (1993) found that GPI-anchored proteins obtained from the erythrocytic stage of Plasmodium falciparum were able to stimulate the synthesis of IL-1 and tumor necrosis factor-α (TNF-α) by murine macrophages. Chow et al. (2010) found that macrophage nitric oxide production and IL-1b mRNA transcription were stimulated strongly after treating chicken macrophages with recombinant SAgS 4, 5 or 12 from E. tenella merozoites. With regard to the role of GPI-anchored surface proteins and cell invasion, Jahn et al. (2009) found that invasion by E. tenella sporozoites was reduced in the presence of monoclonal antibody (2H10E3) that interacts with a well characterized GPI-anchored SAg (EtSAG1). The important roles of GPI-anchored SAgS can play in the interaction between parasites and host cells (attachment, invasion and immune activation or stimulation) was the impetus to investigate GPI-anchored SAgS of E. maxima GS and E. maxima M6 as described in the present work.

Transcriptome sequencing at the depth achieved in the present study has not been accomplished previously from the sporogonic development of E. maxima (although limited EST or cDNA data from E. maxima is publically available from other life cycle stages, e.g. Schwarz et al. 2010). Of more significance is the comparative nature of the transcriptome data obtained in the present study. For the first time, two strains of the same species of Eimeria (GS and M6 strains of E. maxima) were compared at the transcription level under standardized conditions of sporulation and excystation (Al-Badri and Barta 2012). This is also the first comprehensive transcriptome data available for the sporogonic development of E. maxima of any strain and sets the standard for future work. Regardless of the lack of information that is available about the genome of E. maxima, ~50% of the studied transcripts indicated a putative biological function such as metabolic and cellular process, locomotion, signalling, surface antigen etc. The result of annotation and mapping of the investigated transcripts in GS and M6 strains demonstrated a high degree of sequence
similarities between these two strains. It was obvious that most of the transcripts identified in GS and M6 strains have the same functions with similar percentages, shown in Table 4.2. The observed similarities in the transcriptome of these two parasites confirm that these two strains are closely related to one another as previously suggested (Barta et al 1998; Ogedengbe et al. 2011).

The similarities between these two strains were also detected in the number of the putative surface antigens that were determined in GS and M6. Not only was the number of expressed SAg transcripts similar between these two parasites; most of the coding regions of SAgs from a single genetic locus were as well. The alignment of these transcripts (and their associated coding sequences) with comparable strains of *Eimeria maxima* such as the Weybridge and Houghton strains also demonstrated high similarities in the putative coding regions. Although the total numbers of putative GPI-anchored surface antigens expressed by *E. maxima* GS and M6 were similar, it was intriguing that some of these GPI-anchored surface antigens were expressed in only one strain. Although there were a few SAgs that were expressed by GS strain but not expressed in the M6 strain, and vice versa, this limited number of uniquely expressed putative surface antigens is the first demonstration of strain-specific expression of SAgs in *Eimeria* species. This remarkable observation suggests a reasonable explanation for the lack of cross-protective immunity between the two parasite strains. These limited antigenic differences in the sporozoites of the Guelph and M6 strains of *E. maxima* might be major determinants governing the lack of immunological cross-reactivity observed between these parasite strains (e.g. Martin et al. 1997; Beattie 1997; Beattie et al. 2001). The molecular basis of such a profound lack of immunological cross-protection between two *E. maxima* strains was unknown; the present demonstration of differential expression of GPI-anchored surface antigens by *E. maxima* GS and M6 strains coupled with demonstrated polymorphisms at positively selected loci might be sufficient to explain the antigenic uniqueness of each strain, as far as the infected chicken is concerned (e.g. Beattie et al. 2001). Perhaps more surprising is the observation that the differentially expressed surface antigens (Type I, see Figure 5.5) were highly conserved at each genetic locus across at least three strains of *E. maxima* (Wey, H and either GS or M6, depending on expression). It is likely then that both *E. maxima* GS and *E. maxima* M6 possess the same gene repertoire as the Houghton and Weybridge strains and
that expression is truly regulated. Perhaps genome sequencing of these regions of the GS and M6 strains of *E. maxima* would be able to discover subtle regulatory sequences upstream from the Type I SAgs that could explain the selective expression of two of these loci in one parasite strain exclusively and expression of only the third locus in the other strain (again, exclusively). In summary, observations on SAg expression during sporogony of two *E. maxima* strains have proven that the two strains of *E. maxima* (M6 and GS) belong to same species but that there were differences in the repertoire of GPI-anchored surface antigens (SAgs) that were expressed by these two parasites. Uniquely expressed SAgs were identified in each one of these two strains and the abundance of expression was also different among SAgs (although usually similar for loci expressed by both species). The immune system of the chicken might be able to distinguish and interact with these two strains as different parasites because of the differences in their SAg coats (GPI-anchored surface antigens).

The uniquely expressed GPI-anchored surface antigens might be used in future to develop a subunit vaccine that includes molecules from both strains. The immunogenicity of the identified GPI-anchored surface antigen needs to be tested *in vivo* first to understand the ability of these SAgs to stimulate the immune system of chickens and generate protective immunity against subsequent challenge with live *E. maxima* parasites. The discovery of uniquely expressed GPI-anchor surface antigens and some shared SAgs under positive selective pressure represents an important step toward developing such a vaccine. The efficiency of subunit vaccines in protecting chickens against infection with *Eimeria* has been tested in several studies using molecules from different stages of the *Eimeria* life cycle. For instance, a DNA vaccine targeting the gametophyte antigen Gam56 from *Eimeria maxima* in chickens was constructed by Xu et al. (2013). The administration of pcDNA-Gam56 vaccine markedly increased the lymphocyte proliferation activity (P<0.05) at day 7 and 14 after the first immunization. These results indicate that a median dosage of a DNA vaccine had good immunogenicity and immune protection effects, and may be used in field applications for coccidiosis control. In another study (Qi et al. 2013), the recombinant proteins of *E. tenella* mature MIC-1 and adhesive domain (vonWillebrand factor type A domain, EtMIC-1-VD) were inoculated to chickens that were later challenged. It was found that the injection of 100 μg EtMIC-1 or EtMIC-1-VD antigen
induced transcription profiles of IL-12 and IFN-γ of the immunized groups and tenfold higher level of IgG was produced in immunized chickens compared to the control group. It was also found that both DNA vaccines containing the SO7 gene and recombinant SO7 protein improved body weight performance and cecal lesions in challenged chickens compared with controls (Song et al. 2013). Moreover, Liu et al. (2013) found that levels of interleukin-2 (IL-2), interferon-γ (IFN-γ), and proportion of CD4(+) and CD8(+) T lymphocytes increased after birds were immunized with the pVAX1-Rho antigen and later challenged with *E. tenella* oocysts, suggesting that the vaccine stimulated both humoral and cell mediated immunity. The immunized birds also demonstrated reduced oocyst excretion, decreased cecal lesion scores and increased body weight gains compared to the non-immunized chickens.
Chapter 7 – Conclusions

Chapter 3

1. The sporulation stage of life cycle in *Eimeria maxima* GS and *Eimeria maxima* M6 was divided into five morphologically distinguishable stages whose abundance peaked at the following times during sporulation: unsporulated oocysts at 0 h; sporoblast anlagen at 18 h; sporoblasts without sporocyst walls at 22 h; and sporocysts without mature sporozoites at 38 h.

2. Kinetics of oocysts sporulation and peak of shedding between these two immunologically distinct strains of *E. maxima* (GS and M6) was similar.

3. Mean total oocyst output determined for each strain demonstrated that the fecundity of the M6 strain of *E. maxima* was roughly twice that of the GS strain when inoculated at the rate of 1,000 infective oocysts per bird.

4. Both strains had a prepatent period of approximately 120 h followed by peak oocyst shedding at 144-150 h post inoculation.

Chapter 4

1. The transcriptome of sporulating oocysts and fully sporulated, excysting sporozoites of *E. maxima* GS and *E. maxima* M6 were sequenced and analyzed using NGS for first time.

2. Blasting 10,000 transcripts obtained from GS and M6 strains against GenBank non redundant nucleotide database (nr) discovered that ~50% of transcripts had blast hits.

3. Functional annotation and mapping of 10,000 transcripts generated from *E. maxima* GS and *E. maxima* M6 discovered that 60% of generated transcripts have metabolic and cellular process.

4. The *E. maxima* GS and *E. maxima* M6 transcripts showed similarities with sequences from other parasites, particularly from the phylum Apicomplexa. Of transcripts from both strains (GS and M6), the largest number (~40%) of transcripts generated high similarity hits with genes of *Toxoplasma gondii* when a Blast search was conducted against public sequence databases.
5. Blast2GO analysis of sequencing data obtained from both strains discovered 19 putative surface antigens in *E. maxima* GS and 18 putative surface antigens in *E. maxima* M6.

**Chapter 5**

1. Glycosylphosphatidylinositol (GPI) anchored surface antigens (SAgs) were discovered in *E. maxima* GS (18 SAgs) and *E. maxima* M6 (18 SAgs).
2. The alignment of the 36 SAgs with genomic sequences of *E. maxima* Houghton strain and *E. maxima* Weybridge strain discovered that these SAgs were aligned with 21 genes.
3. The alignment of the 36 discovered SAgs discovered that 32 SAgs belong to 4 multicopy gene families and 4 SAgs belong to 2 single copy gene families.
4. The comparison of the identified GPI anchored SAgs between the two studied strains discovered that 3 GPI anchored SAgs were expressed in GS strain only and 3 GPI anchored SAgs were expressed in M6 strain only.
5. The abundance of expression of unique SAgs was lower than the abundance of expression of SAgs that are expressed in both GS and M6 strains.
6. Analysis of SNPs among orthologs shared by up to four strains of *Eimeria maxima* demonstrated that most SAgs were not under positive selective pressure (Types B, C, D, E, J, H); however, shared SAgs of Type A (6 of 8 loci) and Type G (1 of 1 locus) showed high $\delta_N/\delta_S$ ratios indicative of positive, likely immunological, selective pressure.
7. Finally, the discovered differences in the abundance and uniqueness in the expressed GPI-anchored surface antigens between *E. maxima* GS and *E. maxima* M6 might explain the lack of cross-protective immunity elicited in chickens by these the two parasite strains.
References


Carle GF, Frank M, Olson MV (1986) Electrophoretic separations of large DNA molecules by periodic inversion of the electric field. Science 232(4746): 65-68


Ding X, Lillehoj HS, Dalloul RA, Min W, Sato T, Yasuda A (2005) In ovo vaccination with the *Eimeria tenella* EtMIC2 gene induces protective immunity against coccidiosis. Vaccine 23(28): 3733-3740


Edgar SA (1955) Sporulation of oocysts at specific temperatures and notes on the prepatent period of several species of avian coccidia. J Parasitol 41: 214-216


112
Rose ME (1967) Immunity to Eimeria brunetti and Eimeria maxima infections in the fowl. Parasitology 57(2): 363-370
Rose ME, Hesketh P (1976) Immunity to coccidiosis: Stages of the life-cycle of Eimeria maxima which induce, and are affected by, the response of the host. Parasitology 73(1): 25-57
Rose ME, Hesketh P (1976) Immunity to coccidiosis: Stages of the life-cycle of Eimeria maxima which induce, and are affected by, the response of the host. Parasitology 73(1): 25-57
Rose ME, Long PL (1962) Immunity to four species of Eimeria in fowls. Immunology 5:79-92


Shirley MW (1975) Enzyme variation in eimeria species of the chicken. Parasitology 71(3): 369-376


### APPENDIX 3.1 Total oocyst counts in the feces of chickens infected with the M6 or GS strains of *Eimeria maxima*

| Strain/Bird No. | 126 h PI | 132 h PI | 138 h PI | 144 h PI | 150 h PI | 156 h PI | 162 h PI | 168 h PI | 174 h PI | 180 h PI | 186 h PI | 192 h PI | 198 h PI | 204 h PI | 210 h PI | 216 h PI |
|----------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| GS/C1          | 49028   | 46354   | 91908   | 569829  | 1585080 | 133200  | 19980   | 75924   | 143856  | 11988   | 65534   | 11188   | 10389   | 0       | 0       | 0       |
| GS/C2          | 11988   | 28771   | 13586   | 1706292 | 1654344 | 599400  | 89910   | 106293  | 1512086 | 63936   | 0       | 8391    | 14918   | 0       | 0       | 0       |
| GS/C3          | 9191    | 88711   | 212587  | 2949048 | 2904426 | 93240   | 12121   | 143856  | 259740  | 13586   | 38361   | 24775   | 21578   | 0       | 0       | 5194    |
| GS/C4          | 3763    | 63270   | 145054  | 4553841 | 2135430 | 246420  | 64003   | 212086  | 281318  | 36763   | 11188   | 0       | 12787   | 0       | 0       | 7592    |
| GS/C5          | 13330   | 57942   | 143856  | 3178018 | 2904426 | 93240   | 12121   | 143856  | 259740  | 13586   | 38361   | 24775   | 21578   | 0       | 0       | 0       |
| Mean           | 17460   | 57009.6 | 121398.2| 2591405.6| 2821842 | 331668  | 41460.4 | 303615.8| 476163.2| 37402.4 | 28611   | 16063.6 | 17528.8 | 0       | 0       | 2557.2  |
| Std Dev        | 18024.45| 22113.35| 73946.753| 1516028.8| 2274275 | 244903.33| 33855.082| 381014.31| 24810.03| 25401.422| 3602.7677| 231 | 1598 | 3602.7677|
| M6/C1          | 19780   | 1698030 | 349250  | 7987204 | 3263400 | 176490  | 652730  | 587412  | 419580  | 18781   | 17182   | 46353   | 32367   | 5328    | 4195    | 8524    |
| M6/C2          | 20512   | 2229777 | 139860  | 1278720 | 5543118 | 143190  | 29970   | 503496  | 36763   | 39960   | 311688  | 9990    | 197002  | 2131    | 1598    | 3862    |
| M6/C3          | 21578   | 153247  | 100899  | 2224173 | 8301024 | 96570   | 58275   | 430788  | 450478  | 19980   | 14385   | 23976   | 12387   | 2930    | 3196    | 2131    |
| M6/C4          | 40759   | 264269  | 294905  | 4967126 | 4209120 | 632700  | 137862  | No sample| 1096902 | 99900   | 0       | 3729    | 175824  | 0       | 7192    | 6926    |
| M6/C5          | 4928    | 314685  | 202197  | 1158640 | 8511480 | 596400  | 37230   | 1389680 | 783216  | 179620  | 284914  | 536263  | 290109  | 4995    | 0       | 8325    |
| Mean           | 21511.4 | 932001.6| 217382.2| 3503213 | 5965628.4| 329070  | 173213.4| 727871  | 557441.8| 71688.2 | 125633.8| 124062.2| 141537.8| 3076.8  | 3236.2  | 5953.6  |
| Std Dev        | 12742.451| 962344.25| 103988.16| 2917483.4| 2371777.2| 262461.8| 243903.27| 445908.68| 401088.25| 68862.145| 158041.36| 231007.58| 117178.73| 2186.903| 2726.0758| 2835.6268|
### APPENDIX 3.2

One hundred viable oocyst counts every 4 hours to study the kinetics of oocyst sporulation of *Eimeria maxima* M6

<table>
<thead>
<tr>
<th>Sporulation Time</th>
<th>Unsporulated</th>
<th>Sporoblast Anlagen</th>
<th>Sporoblasts without sporocyst walls</th>
<th>Sporocysts without mature sporozoites</th>
<th>Sporocysts with mature sporozoites (RB visible)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>99</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>36</td>
<td>60</td>
<td>40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>5</td>
<td>15</td>
<td>80</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>28</td>
<td>0</td>
<td>10</td>
<td>42</td>
<td>48</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>0</td>
<td>5</td>
<td>22</td>
<td>73</td>
<td>0</td>
</tr>
<tr>
<td>36</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>89</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>2</td>
<td>8</td>
<td>90</td>
<td>0</td>
</tr>
<tr>
<td>44</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>48</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td>52</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>56</td>
<td>44</td>
</tr>
<tr>
<td>56</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>29</td>
<td>71</td>
</tr>
<tr>
<td>60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>64</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>68</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>72</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>138</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>
APPENDIX 3.3 One hundred viable oocyst counts every 4 hours to study the kinetics of oocyst sporulation of *Eimeria maxima* GS

<table>
<thead>
<tr>
<th>Time of sporulation</th>
<th>Unsporulated</th>
<th>Sporoblast Anlagen</th>
<th>Sporoblasts without sporocyst walls</th>
<th>Sporocysts without mature sporozoites</th>
<th>Sporocysts with mature sporozoites (RB visible)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>99</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>99</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>98</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>40</td>
<td>55</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>16</td>
<td>6</td>
<td>88</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>28</td>
<td>0</td>
<td>16</td>
<td>31</td>
<td>53</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>0</td>
<td>7</td>
<td>26</td>
<td>78</td>
<td>0</td>
</tr>
<tr>
<td>36</td>
<td>0</td>
<td>1</td>
<td>25</td>
<td>74</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>87</td>
<td>0</td>
</tr>
<tr>
<td>44</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>98</td>
<td>0</td>
</tr>
<tr>
<td>48</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>97</td>
<td>2</td>
</tr>
<tr>
<td>52</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>61</td>
<td>39</td>
</tr>
<tr>
<td>56</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>32</td>
<td>68</td>
</tr>
<tr>
<td>60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>64</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>68</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>72</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>138</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>
APPENDIX 4.1 MOPS running buffer preparation

10X MOPS Running Buffer (1 L)

Ingredients:
41.9 g of MOPS3-(N-morpholino) propanesulfonic acid
6.8 g Sodium acetate
20 ml of 0.5 M EDTA
900 ml distilled water

Instructions: Combine the above and then adjust the pH of the buffer to 7.0 with NaOH before bringing up the volume to 1 litre with distilled water. The buffer was then autoclaved on liquid cycle for 1 h prior to use.

Saline A

Ingredients:
8.00g NaCl
0.40g KCl
0.35g NaHCO₃
1.00g dextrose, and
1.00 distal water

Instructions:

The ingredients above were mixed together in 2 litters bottle and then filter sterilized prior to use.
### APPENDIX 4.2 Example of Blast2GO Mapping of *Eimeria maxima* M6 transcripts

<table>
<thead>
<tr>
<th>#</th>
<th>sequence name</th>
<th>seq description</th>
<th>length</th>
<th>#hits</th>
<th>min eValue</th>
<th>min mean</th>
<th>#GOs</th>
<th>GO IDs</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Contig_130_M6_...</td>
<td>tubulin beta</td>
<td>1607 20</td>
<td>0.00</td>
<td>0.03%</td>
<td>8</td>
<td>GO:0006974, GO:0005198, GO:0005737, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Contig_2765_M6_...</td>
<td>hypothetical protein</td>
<td>940 3</td>
<td>0.75</td>
<td>B.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Contig_490_M6_...</td>
<td>alpha tubulin</td>
<td>1723 20</td>
<td>0.00</td>
<td>0.75%</td>
<td>8</td>
<td>GO:0006974, GO:0005198, GO:0005737, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Contig_282_M6_...</td>
<td>histone H3</td>
<td>1105 20</td>
<td>0.08</td>
<td>0.63</td>
<td>7</td>
<td>GO:0000735, GO:0001549, GO:0003677, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Contig_125_M6_...</td>
<td>histone H3</td>
<td>1157 20</td>
<td>0.08</td>
<td>0.73</td>
<td>7</td>
<td>GO:0000735, GO:0001549, GO:0003677, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Contig_48_M6_...</td>
<td>histone H3</td>
<td>2855 20</td>
<td>0.08</td>
<td>0.55</td>
<td>7</td>
<td>GO:0000735, GO:0001549, GO:0003677, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Contig_29_M6_...</td>
<td>histone H3</td>
<td>2802 20</td>
<td>0.08</td>
<td>0.53</td>
<td>7</td>
<td>GO:0000735, GO:0001549, GO:0003677, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Contig_43_M6_...</td>
<td>histone H3</td>
<td>2855 20</td>
<td>0.08</td>
<td>0.55</td>
<td>7</td>
<td>GO:0000735, GO:0001549, GO:0003677, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Contig_29_M6_...</td>
<td>histone H3</td>
<td>2802 20</td>
<td>0.08</td>
<td>0.53</td>
<td>7</td>
<td>GO:0000735, GO:0001549, GO:0003677, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Contig_29_M6_...</td>
<td>histone H3</td>
<td>2802 20</td>
<td>0.08</td>
<td>0.53</td>
<td>7</td>
<td>GO:0000735, GO:0001549, GO:0003677, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Contig_850_M6_...</td>
<td>same threonine protein</td>
<td>2860 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>21</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Contig_504_M6_...</td>
<td>same threonine protein</td>
<td>2916 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Contig_902_M6_...</td>
<td>26s protease regulatory subunit 6a</td>
<td>2874 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Contig_902_M6_...</td>
<td>26s protease regulatory subunit 6a</td>
<td>2874 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Contig_2224_M6_...</td>
<td>26s protease regulatory subunit 6a</td>
<td>2633 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Contig_2424_M6_...</td>
<td>26s protease regulatory subunit 6a</td>
<td>2633 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Contig_2424_M6_...</td>
<td>26s protease regulatory subunit 6a</td>
<td>2633 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Contig_2424_M6_...</td>
<td>26s protease regulatory subunit 6a</td>
<td>2633 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Contig_16_M6_...</td>
<td>hypoxia-inducible factor</td>
<td>1792 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Contig_16_M6_...</td>
<td>hypoxia-inducible factor</td>
<td>1792 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Contig_2424_M6_...</td>
<td>hypoxia-inducible factor</td>
<td>1792 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Contig_2424_M6_...</td>
<td>hypoxia-inducible factor</td>
<td>1792 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Contig_470_M6_...</td>
<td>heat shock protein 70</td>
<td>1857 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Contig_2424_M6_...</td>
<td>heat shock protein 70</td>
<td>1857 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Contig_470_M6_...</td>
<td>heat shock protein 70</td>
<td>1857 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Contig_2424_M6_...</td>
<td>heat shock protein 70</td>
<td>1857 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Contig_2424_M6_...</td>
<td>heat shock protein 70</td>
<td>1857 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Contig_2424_M6_...</td>
<td>heat shock protein 70</td>
<td>1857 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Contig_2424_M6_...</td>
<td>heat shock protein 70</td>
<td>1857 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Contig_2424_M6_...</td>
<td>heat shock protein 70</td>
<td>1857 20</td>
<td>0.00</td>
<td>0.50%</td>
<td>24</td>
<td>GO:0004597, GO:0001543, GO:0001906, GO:0006926</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>