UNIVERSITI PUTRA MALAYSIA

HIGH-THROUGHPUT SEQUENCING AND ANALYSIS OF'CHROMOSOME 1 OF EIMERL4 TENELLA

LING KING HWA

FPSK(M) 2005 10

HIGH-THROUGHPUT SEQUENCING AND ANALYSIS OF CHROMOSOME 1 OF *EIMERIA TENELLA*

By

LING KING HWA

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfillment of the Requirements for the Degree of Master of Science

June 2005

This thesis is dedicated to my beloved Pike See

for her unwavering support, understanding, strength,

love, constant encouragement and for instilling me

with a belief in my own abilities.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Science

HIGH-THROUGHPUT SEQUENCING AND ANALYSIS *OF* CHROMOSOME 1 OF *EIMERIA TENELLA*

By

LING KING HWA

June 2005

Chairman : Associate Professor Rozita Rosli, Ph.D.

Faculty : Medicine and Health Sciences

Eimeria tenella is one of seven *Eimeria* species that causes avian coccidiosis. It is also highly pathogenic and is one of the three most common species occurring in the field. High-throughput sequencing of the 1.05Mb chromosomes one of *Eimeria tenella* (Houghton strain) revealed genomic information which may be useful in the discovery of genes such as those involved in drug resistance, cellular regulation and integration, and mechanisms of invasion.

High-throughput and random chromosomal shotgun sequencing resulted in 61 unfinished contigs representing full shotgun state of chromosome one nucleotide sequences which are ready into the finishing phase. Out of these contigs, 57 of them were arranged into 11 scaffolds whereas 4 contigs remain unordered. All the contigs represent 86.9% or 9.29-fold coverage of chromosome one. The quality of the assembly is assured with 94.1% of consistent paired reads with only 5824.3 (0.5%) errors expected. In addition, contiguity of the assembly vastly improved with the integrated BAC-end sequences and HAPPY map markers. Consensus level assessment showed 99.2% of the unfinished chromosomal sequence has expected error rate less than 1 per 10,000 bases (PHRAP score > 40) and only 7.6% of them need further polishing.

The GC content of chromosome one is 49.35% and long-ranged excursions from its mean are found prominently in three regions whereas chromosomal wide GC fluctuations ranged from 35% to 60% at a 12kb window length analysis. GC skews were found to be correlated with the repeats rich regions of the chromosome. Telomeric sequence at both ends of the chromosome is derived as TTTAGGG / CCCTAAA with undefined real length. A centromeric like region with approximately 1,453bp was found in chromosome one with 81.3% AT composition. Chromosome one is expected to bear at least 25.3% of repetitive elements with the most prominent tandem repeat, TGC, which are distributed throughout the chromosome. The longest minisatellite, *mst1*, is 3,624bp in length and occurs as a single stretch in the chromosome. Besides that, there are a few under-characterized interspersed repeats such as LINE and DNA transposons which were found in the chromosome and preliminary homology-based gene survey demonstrated the

possibility of LTR elements in SC11. Both the GC skews and distribution of repetitive elements divide the chromosome into 7 prominent regions.

Alignments with non-redundant and EST databases during gene survey gave a coarse estimation of coding densities of chromosome one at 1 CDS per 1000bp which also corresponded to 12.6% as coding and 87.4% as non-coding. Careful inspection on the distribution revealed that the coding sequences are centrically arranged within the chromosome. GC composition (53.9%) is higher in coding sequences compared to non-coding sequences (48.6%). The number of genes embedded in chromosome one is unknown until further laboratory investigations are carried out. Some of the significant hits may reflect the presence of the genes in chromosome one such as previously characterized LPMC-61 antigen, elongation factor Tu, proteophosphoglycan, proteases, and AAA ATPase family proteins that are involved in the parasite's mobility, parasite-host interaction and possibly invasion.

However, *in silico* gene prediction using a homology-based technique identified three full length genes, phosphatidylinositol-4-phosphate 5-kinase (PIP5K), glucose-6-phosphate isomerase (PGI) and malate:quinone oxidoreductase (MQO). These genes served as gene models and provided early information regarding the intron, exon and splicing sites. The average exon and intron sizes were predicted as 118.5bp and 535.3bp, respectively. The most commonly utilized splice pairs is AG...GT. Chromosome one nucleotide sequences have been deposited in the data depository of the Interim Laboratory of National Institute for the Genomics and Molecular Biology, BIOVALLEY-UKM, Bangi, Malaysia.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains

PENJUJUKAN BERSKALA BESAR DAN ANALISIS KROMOSOM SATU *EIMERIA TENELLA*

Oleh

LING KING HWA

Jun 2005

Pengerusi : Profesor Madya Dr. Rozita Rosli, Ph.D.

Fakulti : Perubatan dan Sains Kesihatan

Eimeria tenella merupakan satu daripada tujuh spesies *Eimeria* yang menyebabkan koksidiosis ayam. Ia juga merupakan spesies yang paling patogenik dan paling kerap ditemui berbanding dengan yang lain. Penjujukan kromosom satu *Eimeria tenella* strain Houghton secara berskala besar akan menghasilkan maklumat genomik untuk penemuan gen-gen yang terlibat dalam kerintangan dadah, integrasi dan pengawalaturan sel serta mekanisme jangkitan.

Penjujukan kromosom berskala besar secara rawak telah menghasilkan 61 kontig yang sedia menjejaki fasa penghabisan. Lima puluh tujuh kontig daripada kontig-kontig ini, telah diaturkan ke dalam 11 kerangka manakala 4 kontig yang lain tidak dapat diaturkan ke dalam mana-mana perancah. Kontig-kontig ini mewakili 86.9% atau 9.29 kali liputan kromosom satu. Pasangan jujukan konsisten yang tinggi (94.1%) dan jangkaan ralat yang rendah (5824.3 atau 0.5%) menjamin kualiti perhimpunan jujukan yang tinggi. Kesinambungan perhimpunan jujukan juga diperbaik melalui integrasi jujukan-jujukan hujung kromosom buatan bakteria dan penanda peta HAPPY. Sebanyak 99.2% daripada jujukan konsensus kromosom satu dijangkakan mempunyai kadar ralat kurang daripada 1 bes dalam setiap 10,000 bes (Skor PHRAP > 40) dan cuma 7.6% daripada jujukan konsensus ini perlu diperbaiki.

Kandungan GC kromosom satu ialah 49.35% dan penyisihan julat besar daripada nilai min didapati ketara di tiga kawasan manakala julat naik-turun kandungan GC secara keseluruhan adalah dari 35% hingga 60% pada analisis gelangsar tingkap bersaiz 12kpb. Sisihan kandungan GC juga didapati berkaitan dengan kawasan kromosom satu yang kaya dengan jujukan ulangan. Jujukan telomer pada kedua-dua hujung kromosom mengandungi jujukan ulangan TTTAGGG / CCCTAAA dengan saiz sebenar belum dapat ditentukan lagi. Kawasan mirip sentromer dengan saiz 1,453pb dan kandungan AT setinggi 81.3% juga ditemui. Kromosom satu dijangkakan mengandungi sekurang-kurangnya 25.3% elemen ulangan. Jujukan ulangan jenis tandem yang paling kerap ditemui ialah TGC dan tertabur sepanjang kromosom satu. Satelit mini yang paling panjang ialah *mst1*, mempunyai saiz 3,624pb dan wujud cuma dalam satu bentangan dalam kromosom. Selain itu, elemen seperti transposon DNA dan elemen nuklear celahan panjang juga ditemui dalam kromosom satu. Peninjauan gen berdasarkan konsep homologi menunjukkan

kewujudan elemen terminal panjang di SC11. Kedua-dua ciri sisihan kandungan GC dan elemen ulangan telah membahagikan kromosom kepada tujuh kawasan yang menonjol.

Penjajaran jujukan pengkodan kromosom satu dengan jujukan-jujukan penanda terungkap dan jujukan-jujukan di pangkalan data '*non-redundant*' semasa peninjauan gen, menganggarkan 1 jujukan pengkodan dalam setiap 1000pb. Ini juga sepadan dengan 12.6% daripada jujukan kromosom satu sebagai jujukan pengkodan dan 87.4% sebagai bukan pengkodan. Pemantauan terperinci menunjukkan bahawa taburan jujukan pengkodan ini diatur ketengah kromosom. Kandungan GC didapati lebih tinggi (53.9%) pada jujukan pengkodan berbanding bukan pengkodan (48.6%). Jumlah sebenar gen yang terkandung dalam kromosom satu masih tidak diketahui sehingga bukti makmal diperolehi. Penjajaran yang ketara menunjukkan kehadiran gen-gen seperti yang mengkodkan antigen LPMC-61, factor pemanjangan Tu, *proteophosphoglycan*, protease dan protein dalam famili AAA ATPase yang kesemuanya terlibat sama ada dalam mobiliti parasit, interaksi hos-parasit dan kemungkinan dalam proses jangkitan.

Ramalan gen secara *in silico* melalui teknik homologi telah menemukan tiga gen bersaiz penuh iaitu, *phosphatidylinositol-4-phosphate 5-kinase* (PIP5K), *glucose-6-phosphate isomerase* (PGI) dan *malate:quinone oxidoreductase* (MQO). Gen-gen ini berfungsi sebagai model gen dan memberi maklumat awal berkaitan intron, ekson dan tapak-tapak penyambatan. Saiz purata ekson dan intron ialah 118.5pb dan 535.3pb masing-masing. Tapak penyambatan yang paling kerap digunakan ialah pasangan AG...GT. Jujukan-jujukan kromosom satu telah disimpan dalam pangkalan data di Makmal Interim untuk Institut Kebangsaan Genomik dan Biologi Molekul BIOVALLEY-UKM, Bangi, Malaysia.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Associate Professor Dr. Rozita Rosli for her guidance and advice. Her support and encouragement make this study a successful one. I thank her again for trusted me so much in this project and I appreciate her 3-year supervision that allowed me to learn in the greatest pace of my life. I also salute and respect her openness during our discussion regarding the project and all the help she provided as a mentor and sometimes even across the boundaries, to ensure a bright future for her students.

I would like to express my deepest gratitude to my second supervisor, Associate Professor Dr. Wan Kiew Lian for his everlasting input during most of the informal and formal discussions. Thanks for all the hospitality and care that he provided during our visit to UK. Without his great concern both financially and personally, I would not be able to expand my experience with renowned scientists at the Institute for Animal Health and Wellcome Trust Sanger Institute, UK.

Thanks are also extended to my third supervisor, Associate Professor Dr. Mariana Nor Shamsudin for her inspiration and suggestions throughout the study. I really appreciate all her efforts to make me think maturely and feel grateful for this piece of genomic- and bioinformatics-based project in the initial state of my study. She always reminds me about planning and implementation of the project. I will always bear in mind her advice that put great emphasis on time management and to graduate promptly according to schedule. Thank you so much Dr.!

I also wish to thank my closest colleagues, Esther and Wai Yan who taught me from scratch and were always the first to contribute towards the completion of chromosome one sequencing project. Not to forget Shu San and Soon Joo, thanks for every finger that you shared throughout the study. Your help during the plasmid and BAC extractions will be always in my mind and down to my heart I wish you guys all the best in all your future undertakings. Special thanks to the lab manager, Mat Noor, for all his patience and guidance in teaching me UNIX commands. Your helping hands are the most valuable and without you I won't be able to sit in front of the SGI machine for such a long time. Not to forget Professor Dr. Rahmah in granting me permission to stay and work in the interim laboratory for NIGMB. Thank you so much.

Recognition also goes to all the labmates in the Laboratory of Molecular Genetics for sharing my joy and tears throughout the time. To Kak Nurma, Kak Nik, Nasir, Thila, Che Pah, Kak Reza, Kak Sherry Sr. and Jr., Siu Tung, Shaban, Lama, Farouk, Kak Su and Radha – I learnt a lot from you and I wish everyone all the best!

Special credits also go to all the international collaborators who never hesitate to disclose their unpublished findings in contributing towards the completion of this work. I would like to express my gratitude to Professor Martin Shirley (UK), Dr. Fiona Tomley (UK), Dr. Paul Dear (UK), Dr. Al Ivens (UK), Dr. Arthur Gruber (Brazil), Dr. Jianping Cai (China) and Ms. Sarah White (UK).

On a personal note, I would like to express gratitude to my beloved housemates, Pak Guan, Emily and Edward for their concerns throughout my stay with them all this while. Thousands of thanks also go to both of my parents and siblings who provide me with the courage to explore into this new field of studies and never doubted on my abilities to complete the whole course for the degree. Finally, special thanks to Pike See for her motivation and considerate heart that make me feel no pressure throughout the study!

Ling King Hwa

12th October 2004

TABLE OF CONTENTS

		Page
DEDICATION		ii
ABSTRACT		iii
ABSTRAK		v
ACKNOWLEDGMENTS		vii
APPROVAL		x
DECLARATION		xii
LIST OF TABLES		xvii
LIST OF FIGURES		xix
LIST OF ABBREVIATIONS		xxii
CHAPTER		
1	**INTRODUCTION**	
1.1	*Eimeria species* and coccidiosis	1
1.2	Problem statement	2
1.3	Significance of study	4
1.4	Objectives	6
2	**LITERATURE REVIEW**	
2.1	*Eimeria tenella*	7
2.1.1	Life cycle of *Eimeria tenella*	10
2.1.2	Pathogenicity of *Eimeria tenella*	14
2.2	Degree of pathogenesis	17
2.2.1	Number of asexual cycles and merozoite produced	17
2.2.2	Dosage of oocysts	17
2.2.3	Region of infection	18
2.2.4	Other factors	19
2.3	Host cell invasion by *Eimeria tenella*	19
2.4	Control of coccidiosis in chicken	24
2.4.1	Prophylactic control with anticoccidial drugs	24
2.4.2	Vaccination	25
2.4.3	Alternative coccidiosis control	30
2.5	*Eimeria tenella* (Houghton strain) as candidate for genome sequencing	31
2.6	Genome of *Eimeria tenella*	33
2.6.1	Nuclear genome	33
2.6.2	Apicoplast or plastid genome	36
2.6.3	Mitochondrial genome	38
2.6.4	Double-stranded RNA segments	38
2.7	Chromosome one of *Eimeria tenella* (Houghton strain)	39

2.8 High-throughput DNA sequencing technologies 41
2.8.1 Strategies for High-throughput sequencing of eukaryotic genomes 42
2.8.2 Random sequencing approach 42
2.8.3 Directed sequencing approach 46
2.9 HAPPY mapping 47
2.10 Type of sequence data generated 49
2.11 Annotating eukaryote genomes 52
2.11.1 *Ab initio* method 52
2.11.2 Homology-based method 54

3 **METHODOLOGY**
3.1 General outline of the study 57
3.2 Phase I: Data generation 62
3.2.1 Chromosome one-specific DNA library 62
3.2.2 Transformation 62
3.2.3 Large-scale bacterial culture preparation 63
3.2.4 Storage of bacterial culture plates 63
3.2.5 Large-scale plasmid DNA preparation 64
3.2.5.1 Manifold techniques 64
3.2.5.2 MWG RoboSmart-384Automated Plasmid Preparation 65
3.2.6 Screening of extracted plasmid 65
3.2.7 Plasmid digestion using *Eco*RI 66
3.2.8 Cycle sequencing 66
3.2.9 Ethanol / Sodium acetate precipitation in 96-well reaction plates 67
3.2.10 Denaturation prior to single-stranded DNA sequencing 68
3.2.11 Automated Sequencing using ABI 3100 and ABI 3730 68
3.2.12 Naming convention 69
3.3 Phase II: Assembly of data 70
3.3.1 Preprocessing of Raw AB1 files 70
3.3.1.1 Base confidence values 70
3.3.1.2 Quality clipping 71
3.3.1.3 Sequencing vector clipping 71
3.3.1.4 Repeat masking 72
3.3.1.5 Contamination screening 72
3.3.2 Initial assembly of the passed shotgun sequences 74
3.3.3 GAP4 Assembly 75
3.3.4 Characterization of assembly database 77
3.3.4.1 Expected errors and error rate 77
3.3.4.2 Paired read consistency 78

3.4 Phase III: Pre-finishing of chromosome one 79
3.4.1 Introduction 79
3.4.2 Ordering of contigs 79
3.4.3 Preparation of Bacterial Artifical Chromosome (BAC) DNA 80
3.4.4 Digestion of BAC clones 81
3.4.5 Pulse field gel electrophoresis for BAC clones sizing 82
3.4.6 Mapping of scaffolds and contigs onto HAPPY map 82
3.5 Phase IV: Data analysis 83
3.5.1 Consensus calculation using confidence values algorithm 83
3.5.2 Quality of consensus 83
3.5.3 GC content 84
3.5.4 Repeat contents 84
3.5.4.1 Simple Sequence Repeats 84
3.5.4.2 Interspersed repeats 84
3.5.5 Sequence alignments 85
3.5.6 Gene survey 86
3.5.7 Annotation of model gene 87
3.5.8 Multiple sequence alignment 87

4 SEQUENCE GENERATION AND ASSEMBLY
4.1 Introduction 88
4.2 Results and discussion 91
4.2.1 Transformation efficiency 91
4.2.2 Plasmid DNA extraction efficiency 94
4.2.3 Consistency of insert size 96
4.2.4 Large scale generation of shotgun sequences 98
4.2.5 Screening of shotgun sequences 99
4.2.6 Overview of the shotgun assembly 103
4.2.7 Analysis of assembly consistency 106
4.2.8 Pre-finishing of chromosome one contigs 112
4.2.9 Data Integration with sequences generated for chromosome two 120
4.2.10 Data integration with assembled WGS consensus 120
4.2.11 Ordering of contigs 124
4.2.12 Mapping of scaffolds and contigs using HAPPY map 128
4.2.13 Consensus quality 132
4.2.13.1 Base confidence calculation 132
4.2.13.2 Base quality calculation 133
4.3 General discussion 137
4.3.1 Assembly of shotgun sequences 137
4.3.2 Pre-finishing of chromosome one sequences 138
4.3.3 Consensus quality 139
4.4 Conclusion 141

5 SEQUENCE ANALYSES
5.1 Introduction 142
5.2 Results and discussion 143
5.2.1 GC Content of the chromosome 143
5.2.2 Telomeric-like regions 151
5.2.3 Centromeric-like region 154
5.2.4 Repetitive elements in the chromosome 155
5.2.5 Overview of gene content 162
5.2.5.1 Antigenic proteins 172
5.2.5.2 Adhesion proteins 174
5.2.5.3 Proteases 175
5.2.5.4 Others enzymes 176
5.2.6 Homology-based gene annotation 179
5.2.6.1 Putative PIPK in *Eimeria tenella* 179
5.2.6.2 Putative PGI in *Eimeria tenella* 183
5.2.6.3 Putative MQO in *Eimeria tenella* 187
5.2.6.4 Intron, exon and splicing sites 191
5.3 Conclusion 193

6 GENERAL DISCUSSION
6.1 Introduction 194
6.2 Advancement of *Eimeria*'s Biology 194
6.2.1 Chromosomal organization and structure 194
6.2.2 Non-coding and new regulatory elements 196
6.3 Molecular targets for vaccine development 197
6.4 Simple tandem repeats (STRs) 198
6.5 Conclusion 198

7 CONCLUSION
6.1 Scientific recapitulation 199
6.2 Achievement of the projected objectives 201
6.3 Limitation of the study 202
6.3 Future direction 202

REFERENCES 204
APPENDICES 230
BIODATA OF THE AUTHOR 248

LIST OF TABLES

Table		Page
1	Characteristics of chicken coccidiosis caused by *Eimeria* species	16
2	Live anticoccidial vaccines for breeders	28
3	Live anticoccidial vaccines for broilers	29
4	Summary of the biological features of the Houghton strain of *Eimeria tenella*	32
5	Sizes of chromosome one in different populations of *Eimeria tenella*	40
6	Total apicompexan ESTs and GSSs from GenBank databases - dbEST and dbGSS	50
7	Gene-finding related Web pages and programs for genome annotation	54
8	BLAST (Basic Local Alignment Search Tools) programs provided by National Centre for Biotechnology Information	56
9	Values of scoring for pairwise alignment, banded search and the stringency and completeness during repeat masking	73
10	Stringency and completeness during PHRAP assembly	75
11	Values of scoring for pairwise alignment and the stringency during GAP4 assembly	76
12	Sequence alignments using BLAST programs	85
13	Transformation efficiency (CFU/ng) of chromosome specific libraries	93
14	Summary of the pre-analysis of shotgun sequences	101
15	Characteristics of assemblies at various cutoffs of contigs size	105
16	List of chromosome one – related contigs from assembled WGS consensus dated 7th January 2003	121
17	Characteristics of assembly databases	123

18	Characteristics of generated scaffolds for chromosome one of *Eimeria tenella*	128
19	Number of HAPPY map marker in each scaffold and unordered contigs	130
20	Distribution of PHRAP-equivalent scores in the consensus sequence of chromosome one	133
21	Consensus quality according to categories of data	135
22	Base composition of chromosome one draft sequence	144
23	GC content of various protozoan parasites	144
24	Top 10 most abundant simple sequence repeats in the chromosome one of *Eimeria tenella*	157
25	Characteristics of the top 5 largest minisatellites of the chromosome one of *Eimeria tenella*	157
26	Sequences that matched significantly (non-redundant) with protein of known, similar or putative functions	166
27	Features of exon-intron structure of putative PIP5K gene of *Eimeria tenella*	180
28	Feature of exon-intron structure of putative PGI gene of *Eimeria tenella*	184
29	Features of exon-intron structure of putative MQO gene of *Eimeria tenella*	188
30	Intron, exon and splicing sites of the predicted putative gene of chromosome one of *Eimeria tenella*	192

LIST OF FIGURES

Figure		Page
1	Rod- or banana-shaped sporozoite of *Eimeria* species	8
2	Morphology of a typical oocyst of *Eimeria* species	9
3	Life cycle of chicken coccidian *Eimeria tenella*	13
4	Molecular karyotypes of different strains of *Eimeria tenella*	35
5	Gene organization of the *Eimeria tenella* plastid genome (~35kb)	37
6	Idealized representation of the whole genome shotgun sequencing strategy	44
7	Idealized representation of the hierarchical shotgun sequencing strategy	45
8	HAPPY mapping	48
9a	The pipeline for Phase I of the study; data generation	58
9b	The pipeline for Phase II of the study; assembly of Data	59
9c	The pipeline for Phase III of the study; pre-finishing of data	60
9d	The pipeline for Phase IV of the study; data analysis	61
10	The 96-well high-throughput gel electrophoresis	95
11	Screening of insert size	97
12	Size distributions of passed sequences	101
13	Example of pre-analyzed sequence	102
14	Number of contigs and chromosome coverage vs. number of shotgun sequences computed for assembly	109
15	The percentage of consistent read pairs and chromosome coverage vs. number of sequences computed	110
16	The percentage of base errors generated and chromosome coverage vs. number of sequences computed	111
17	Cleave-like distributions of the paired reads	114

18	Readings with inconsistent size of template	115
19	Misasembled spanning read pairs	116
20	Inconsistent reads after quality re-clipping	117
21	Hairpin loop regions	118
22	Misassembled region showing two distinct sets of quality reads	119
23	Size of selected chromosome one-related BAC clones	125
24	The relative location of the sized BAC clones in the assembly	126
25	The arrangement of contigs according to paired reads (both local and WTSI BAC-end reads)	127
26	HAPPY mapping for chromosome one of *Eimeria tenella* (Houghton strain)	131
27	Consistency display for the chromosome one consensus sequence	136
28	GC composition of chromosome one sequence at 200bp window size	145
29	The GC content with 2.5 standard deviation cutoff at 200bp window size	146
30	Histogram of GC content at 12kb window size in the full shotgun sequence of chromosome one	149
31	Regional average GC content of chromosome one of *Eimeria tenella*	150
32	Distribution of the telomeric-like repeat, TTTAGGG in chromosome one of *Eimeria tenella*	153
33	The centromeric-like region in chromosome one	154
34	Long stretches of repeats in three distinct regions of chromosome one	160
35	Regional average coding densities	171
36	Artemis representation of putative PIP5K gene of *Eimeria tenella*	181
37	Multiple sequence alignment of PIP5K protein sequences	182

38 Artemis representation of putative PGI gene of *Eimeria tenella* 185

39 Multiple sequence alignment of PGI protein sequences 186

40 Artemis representation of putative MQO gene of *Eimeria tenella* 189

41 Multiple sequence alignment of MQO protein sequences 190

LIST OF ABBREVIATIONS

Ω	Ohm
µl	Microliter
A	Adenine
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bp	Base pair
C	Cytosine
cDNA	Complementary deoxyribonucleic acid
ddNTP	Dideoxyribonucleotide triphosphate
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
EST	Expressed sequence tag
F	Faraday unit
G	Guanine
h	Hour
HSP	High Scoring Segments
IPTG	Isopropyl-beta-D-thiogalactopyranoside
Kb	Kilobase
kV	Kilovolt
LB	Luria-Bertani
LINE	Long interspersed nuclear element
LTR	Long terminal repeat
mA	Milliampere

Mb	Megabase
mg	Milligram
ml	Milliliter
MQO	Malate:quinone oxidoreductase
ng	Nanogram
ORF	Open reading frame
PCR	Polymerase chain reaction
PFGE	Pulse-Field Gel Electrophoresis
PGI	Glucose-6-phosphate isomerase
PIP5K	Phosphatidylinositol-4-phosphate 5-kinase
POP	Performance Optimized Polymer
RNA	Ribonucleic acid
rpm	revolution per minute
T	Thymine
U	Unit enzyme
X-gal	5-bromo-4-cloro-3-indolil-beta-D-galatopiranocid
Hg	Mercury

CHAPTER 1

INTRODUCTION

1.1 *Eimeria species* and coccidiosis

Eimeria species belong to the phylum Apicomplexa and is responsible for the diseases of coccidiosis in intensively reared livestock such as poultry, cattle and sheep. Other members of this phylum include *Plasmodium falciparum*, the causative agent of malaria; *Toxoplasma gondii*, an opportunistic pathogen in immunocompromised individuals; and *Cryptosporidium parvum*, an animal parasite as well as an opportunistic pathogen of humans. These parasites are characterized by the presence of apical complex structures or organelles such as conoid, apical polar ring, micronemes, rhoptries and dense granules (Chobotar and Scholtyseck, 1982). Apicomplexans are obligate intracellular parasites that require host cells in which they can invade and replicate to ensure their survival.

Eimeria species from chickens are the most important parasitic pathogens of poultry. There are seven *Eimeria species* that infect chickens; *Eimeria tenella*, *Eimeria necatrix*, *Eimeria acervulina*, *Eimeria brunetti*, *Eimeria maxima*, *Eimeria mitis* and *Eimeria praecox* (Marquardt *et al.*, 2000). All the seven *Eimeria species* infect the intestinal epithelial lining of chicken at only specific locations. In an intensively reared flock of chickens, coccidiosis is always due to the infections of more than one species of *Eimeria* (Williams *et al.*, 1996).