DNA Barcoding of Apicomplexa: Mitochondrial Evolution across the Phylum

by

Mosunmola Elina Ogedengbe

A Thesis
presented to
The University of Guelph

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

Guelph, Ontario, Canada

© Mosunmola E. Ogedengbe, August, 2015
ABSTRACT

DNA Barcoding of Apicomplexa: Mitochondrial Evolution across the Phylum

Mosunmola Elina Ogedengbe
University of Guelph, 2015

Advisor: Dr. J.R. Barta

The parasitic phylum Apicomplexa contains thousands of species identified historically using morphology, morphometrics and host/tissue associations. Diagnostic stages of apicomplexan parasites are difficult to identify morphologically; thus, classical methods were supplemented with molecular techniques. Broadly applicable molecular markers that could support species identification and phylogenetic studies were sorely needed. “DNA barcoding” utilizes short, standardized DNA sequence obtained from the mitochondrial (mt) cytochrome c oxidase subunit I (COI) gene. Lack of PCR primers suitable for apicomplexan parasites limited mt COI use. Research reported herein tested utility of mt COI sequences for identification and phylogenetics of tissue (i.e. Sarcocystidae), eimeriid (e.g. Eimeriidae, Lankesterellidae) and adeleid coccidia (i.e. Hepatozoon and other haemogregarines).

Degenerate primers of broad specificities generated mt COI DNA barcodes that, with few exceptions (notably Cystoisospora spp.), identified any enteric (e.g. Isospora, Cyclospora, Eimeria) or tissue (e.g. Toxoplasma, Hammondia, Sarcocystis) coccidium. The mt COI locus is recommended as a suitable DNA barcoding target for coccidian parasites of veterinary and zoonotic importance.

Complete mt genomes of >20 eimeriid and adeleid coccidia were sequenced. More 17 species representing 6 genera of eimeriid (Eimeria, Isospora, Caryospora, Cyclospora, Lankesterella) and adeleid coccidia (Hepatozoon catesbianae) were sequenced. Eimeriid mt genomes (~ 6,000 bases) are predominantly circular-mapping possessing 3 protein-coding genes (CDS) and rDNA fragments in conserved order; only Cyclospora cayetanensis, a human pathogen, had a linear mt genome possessing a
5′-terminal hairpin and 3′-terminus with telomere-like repetitive elements. Gene orders and directions were largely conserved within, but divergent between, major groups of apicomplexan parasites.

Phylogenetic analyses using mt CDS sequences (i.e. partial mt COI or concatenated CDS sequences) with or without nuclear 18S rDNA sequences suggest combined nu 18S rDNA and mt COI sequences (analysed using partitioned data with suitable nucleotide substitution models for each partition) provide robust evolutionary hypotheses for these parasites. Taxonomic solutions addressing paraphylies identified in phylogenetic analyses were proposed.

The mt genomes of Apicomplexa possess divergent structures and gene arrangements. Mitochondrial COI sequences were concluded to be better species-level markers than the widely used nu 18S rDNA locus; mt COI sequences had utility in both diagnostic and phylogenetic studies.
ACKNOWLEDGEMENTS

I wish to express my gratitude to advisor, Dr John Robert Barta for his continued guidance, insights and support for this work.

I would also like to thank my graduate committee members Drs Scott Weese Tara Gariepy, and Ryan Gregory, for their support, constant interest and suggestions to this Thesis.

My appreciation also goes to the external examiner Dr. Mark C. Jenkins from the US Department of Agriculture for his contributions as external examiner to the completion of my thesis.

I wish to express my sincere thanks to Julia Whale for her support and assistance, Drs H. David Chapman, Bill Chobotar, Chunlei Su, Donald Martin, Bob Rosenthal, Scott Weese and Yvonne Qvarnstrom for their kind provision of isolates and DNA of various coccidia used for this study.

I also appreciate Drs Robert Jacobs and Geoff Wood and Ms Donna Kangas for all the help, counsel and encouragements that they so often provide to me during this study.

I thank the Ontario Veterinary College, National Veterinary Research Institute, Vom Nigeria, the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA), the National Science and Engineering Research Council of Canada (NSERC), for funding this study.

I wish to thank my husband, Joseph Dairo, for standing by and for me through out while constantly catering for himself and the family when I am too busy in the lab to do so. My children and family: Femi, Seyi, Moji “Jisola” and Ken, Gbubemi “Gbubu” and Chene, Eweyeju “Ewe” and Doo, Femi, Alero “Alegege” and Michael, Bukki “Abuke” and Abel. Dr Iyabo, Ms Adeola, Rotimi and Tope, Olumide, Segun and Pat, Olu Ogedengbe, Bola Umar, Mr and Mrs Ezekiel Laka and my father-in-law Pa Samuel Ogedengbe. My friends Bola, Yewande “Coccus”, Ehis, Bro John Ighogboja and Venerable and Mrs Luka Musa. I thank you all for being there always for me, for your enduring supports, prayers and constant encouragements. Wish I could share this degree with you all.

Lastly, I thank the Lord God Almighty for His grace and provisions through it all.
# TABLE OF CONTENTS

ABSTRACT ........................................................................................................................................ i
ACKNOWLEDGEMENTS ........................................................................................................................ iv
TABLE OF CONTENTS .......................................................................................................................... v
LIST OF FIGURES .................................................................................................................................... ix
LIST OF TABLES ....................................................................................................................................... xi
LIST OF APPENDICES ........................................................................................................................... xii
DECLARATION OF WORK PERFORMED ............................................................................................... xiii
1. INTRODUCTION .................................................................................................................................. 1
2. LITERATURE REVIEW ......................................................................................................................... 6
  2.1. BIOLOGY OF THE PHYLUM APICOMPLEXA ................................................................................. 6
  2.2. CLASSIFICATIONS OF SUPER GROUPS ......................................................................................... 10
  2.3. IDENTIFICATION METHODS AND CHALLENGES ....................................................................... 16
    2.3.1. Morphological Identification .................................................................................................. 16
    2.3.2. Molecular Identification, Differentiation & Phylogenetics ...................................................... 18
  2.4. PHYLOGENETIC RELATIONSHIPS IN THE APICOMPLEXA ......................................................... 21
  2.5. DNA BARCODING AND MITOCHONDRIAL GENOME SEQUENCING ......................................... 27
    2.5.1. Mitochondrial Origin And Evolution ...................................................................................... 27
    2.5.2. Apicomplexan Mitochondrial Genomes .................................................................................. 28
    2.5.3. The Barcode of Life Initiative ................................................................................................ 33
    2.5.4. Gene of Choice for Dna Barcoding ......................................................................................... 36
    2.5.5. COI Universal Primers ............................................................................................................ 38
    2.5.6. COI Sequences in the Apicomplexa ......................................................................................... 39
  2.6. PROBLEMS AND CONCLUDING REMARKS ............................................................................... 41
3. RESEARCH OBJECTIVES .................................................................................................................. 46
4. RESEARCH APPROACHES ................................................................................................................ 47
5. LIMITATIONS ...................................................................................................................................... 48
  5.1. CRYPTOSPORIDIUM: AN ALTERNATE BARCODE TARGET REQUIRED ....................................... 48
  5.2. SMALL SAMPLE SIZES ............................................................................................................... 48
  5.3. TYPE SPECIMENS ....................................................................................................................... 49
6. SEQUENCING THE COMPLETE MITOCHONDRIAL GENOME OF EIMERIA MITIS STRAIN USDA 50 (APICOMPLEXA: EIMERIIDAE) SUGGESTS CONSERVED START POSITIONS FOR mtCOI-AND mtCOIII-CODING REGIONS ............................................................................................................. 50
  6.1. ABSTRACT ....................................................................................................................................... 50
  6.2. INTRODUCTION ........................................................................................................................... 50
6.3. MATERIALS AND METHODS .............................................................. 52
  6.3.1. Parasites and DNA Extraction .................................................. 52
  6.3.2. PCR Amplification .............................................................. 53
  6.3.4. Cloning ................................................................................. 54
  6.3.5. Sequencing ........................................................................... 54
  6.3.6. Sequence Data Assembly ......................................................... 54
6.4. RESULTS AND DISCUSSION ......................................................... 56
  6.4.2 Start codon determinations for COI, COIII and CytB. ..................... 57
  6.4.3. Comparison of cloned versus directly sequenced PCR products .... 58
6.5. DISCUSSION ................................................................................. 59
7. COMPLETE MITOCHONDRIAL GENOME SEQUENCES FROM
   FIVE EIMERIA SPECIES (APICOMPLEX; COCCIDIA; EIMERIIAE)
   INFECTION DOMESTIC TURKEYS ....................................................... 65
7.1. ABSTRACT .................................................................................. 65
7.2. INTRODUCTION ........................................................................ 65
7.3. METHODS .................................................................................. 66
  7.3.1. Parasites ................................................................................. 66
  7.3.2. DNA extraction and long PCR amplification .............................. 67
  7.3.3. Sequencing ........................................................................... 68
  7.3.4. Sequence data assembly and analysis ....................................... 68
  7.3.5. Phylogenetic analyses ......................................................... 69
7.4. RESULTS .................................................................................... 71
  7.4.1. Six mt genomes from five Eimeria species infecting turkeys ........ 71
  7.4.2. Genome organization .......................................................... 73
  7.4.3. Start codon determinations for COI, COIII and CytB .................. 74
  7.4.4. Phylogenetic analyses ........................................................... 75
7.5. DISCUSSION ............................................................................. 79
7.6. CONCLUSIONS ....................................................................... 81
8. A LINEAR MITOCHONDRIAL GENOME OF CYCLOSPORA
   CAYETANENSIS (EIMERIIDAE, EUCOCCIDIORIDA, COCCIDIASINA,
   APICOMPLEXA) SUGGESTS THE ANCESTRAL START POSITION
   WITHIN MITOCHONDRIAL GENOMES OF EIMERIID COCCIDIA .......... 83
8.1. ABSTRACT ................................................................................ 83
8.2. INTRODUCTION ..................................................................... 83
8.3. MATERIALS AND METHODS .......................................................... 84
8.4. RESULTS AND DISCUSSION .......................................................... 87
8.5. ACKNOWLEDGEMENTS ................................................................. 93
9. THE COMPLETE MITOCHONDRIAL GENOME OF CARYOSPORA
   BIGENETICA (EIMERIIDAE, EUCOCCIDIORIDA, COCCIDIASINA,
   APICOMPLEXA) ........................................................................... 94
9.1. ABSTRACT ............................................................................ 94
9.2. INTRODUCTION ..................................................................... 94
LIST OF FIGURES

FIGURE 2.1. Life cycles of apicomplexan parasites ................................................. 8
FIGURE 2.2. A view of eukaryote phylogeny. .......................................................... 12
FIGURE 2.3. Apical complex organelles ................................................................. 12
FIGURE 2.4. Hypothetical evolutionary tree of the Apicomplexa ......................... 13
FIGURE 2.5. Evolutionary relationships among major groups ............................. 25
FIGURE 2.6. Circular structure of the mitochondrial genome for Plasmodium gallinaceum .......................................................... 30
FIGURE 2.7. Diversity in the linear mitochondrial genomes of Babesia gibsoni, Theileria orientalis, Theileria equi and Theileria parva ....................... 31
FIGURE 2.8. Diagrammatic representation of a single mitochondrial genome \‘unit\’ of Eimeria species ...................................................................................... 31
FIGURE 6.1. Schematic map for three Eimeria mits linear mitochondrial genome sequences .......................................................................................... 57
FIGURE 6.2. Possible COIII start codon positions for 11 Eimeria species mitochondrial sequences .......................................................... 61
FIGURE 6.3. Possible COI start codon positions for 11 Eimeria species mitochondrial sequences .......................................................... 62
FIGURE 7.1. Mitochondrial genome organization of five Eimeria species \(\text{Apicomplexa; Eimeriidae}\) infecting turkeys .................................................. 73
FIGURE 7.2. Bayesian inference and maximum likelihood phylogenetic reconstructions using mitochondrial genome sequences of 16 Eimeria species .......................................................... 77
FIGURE 7.3. Maximum parsimony phylogenetic reconstruction using mitochondrial genome sequences of 16 Eimeria species ........................................ 78
FIGURE 8.1. Aligned 5\’ - and 3\’-termini of the Cyclospora cayetanensis mitochondrial genome aligned with the homologous regions in the mt genomes of other eimeriid coccidia ........................................................................ 89
FIGURE 9.1. Caryospora bigenetica \(\text{Apicomplexa; Eimeriidae}\) from a Massasauga rattlesnake ........................................................................ 96
FIGURE 11.1. Sporulated oocysts of two tissue coccidian genera ............................ 113
FIGURE 11.2. Phylogeny of eimeriid coccidia \(\text{Eimeriidae and close relatives}\) and isosporoid coccidia \(\text{Sarcocystidae}\) based on the 18S rDNA dataset .......................................................... 117
FIGURE 11.3. Enlarged portion of the 18S rDNA sequence based tree .................. 118
FIGURE 11.4. Phylogeny of isosporoid coccidia based on mt COI ..................... 119
FIGURE 11.5. Phylogeny of isosporoid coccidia based on the concatenated dataset \(\text{18S rDNA and COI sequences}\) .......................................................... 120
FIGURE 12.1. Phylogenetic tree based on the concatenated 18S rDNA and COI sequence dataset .......................................................... 138
FIGURE 12.2. Phylogenetic tree based on the COI sequence dataset ................. 140
FIGURE 12.3. Phylogenetic tree generated from the 18S rDNA dataset............. 142

FIGURE 13.1. Illustrates the screenshot of Mauve alignment for 22 apicomplexan parasites mt genomes ................................................................. 157

FIGURE 13.2. Schematic illustration of gene orders, rearrangements and inversions in four piroplasms ................................................................. 158

FIGURE 13.3. Variable rearrangements and gene orders within a single piroplasm species, Babesia microti ................................................................. 158

FIGURE 13.4. Pairwise comparisons of gene rearrangements between Eimeria and Leucocytozoon, Hepatozoon or Babesia species .................................. 159

FIGURE 13.5. Phylogenetic tree based on concatenated CDS dataset. ............... 163

FIGURE 13.6. Phylogenetic tree based on concatenated amino acid translation dataset ................................................................. 163

FIGURE 14.1. Unrooted BI phylogenetic tree representing relationships among major apicomplexan genera ................................................................. 179
LIST OF TABLES

TABLE 2.1. Taxonomic classification of the phylum Apicomplexa with some exemplar genera and species (after Levine, 1988). .................................. 15
TABLE 2.2. Molecular targets used with various groups of Apicomplexa .......... 19
TABLE 6.1. Sequencing primers and their locations in the mitochondrial genome of *Eimeria mitis* USDA 50 Clone 24 (KC409031)............................... 54
TABLE 6.2. Pairwise sequence identity and number of nucleotide differences among four complete mitochondrial genome sequences of *Eimeria mitis*. .......... 56
TABLE 7.1. Base composition of whole mitochondrial genomes from five *Eimeria* species infecting turkeys................................................................. 72
TABLE 7.2. Percent pairwise sequence identities and number of nucleotide differences for six mitochondrial genomes from five *Eimeria* species that infect turkeys .................................................................................. 72
TABLE 8.1. Amplification primers used to obtain the mitochondrial genome of *Cyclospora cayetanensis* and confirm its linear nature. ....................... 88
TABLE 10.1. Protein coding genes and fragmented rDNA’s found in the complete mitochondrial genome of an *Isospora* sp. (Eimeriidae, Apicomplexa) causing systemic coccidiosis in canaries (*Serinus canaria* Linn). .................................................................................................. 99
TABLE 11.1. Mitochondrial COI and nuclear 18S rDNA amplification primer combinations, anneal temperatures and expected product sizes........... 108
TABLE 11.2. Oocyst and sporocyst dimensions for parasites obtained from fecal samples. ......................................................................................... 113
TABLE 11.3. COI and 18S rDNA sequences used in present study ................. 114
TABLE 13.1. Mitochondrial genome sequences used in phylogenetic analyses... 152
LIST OF APPENDICES

APPENDIX 1: Amino acids encoded by four mitochondrial genomes obtained from various *Eimeria mitis* strains .......................................................... 219

APPENDIX 2: Chapter 7 - Supplementary figures ........................................ 220

APPENDIX 3: List of sequences, strains and accession numbers used in Chapter 12 .......................................................................................... 222

APPENDIX 4: List of all new sequences submitted to GenBank (not used in thesis chapters) ................................................................. 225

APPENDIX 5: List of all new sequences used in the chapters of the thesis only, submitted to GenBank .......................................................... 226

APPENDIX 6: Divergent nuclear 18S rDNA paralogs in a turkey coccidium, *Eimeria meleagrinis*, complicate molecular systematics and identification.... 229

APPENDIX 7: Biological re-description of a genetically-typed, single oocyst line of the turkey coccidium, *Eimeria meleagrinis* Tyzzer 1929 .................... 237

APPENDIX 8: The complete mitochondrial genome sequence of *Hepatozoon catesbiana* (Apicomplexa; Coccidia; Adeleorina), a blood parasite of the green frog, *Lithobates (Rana) clamitans* .................................................. 250

APPENDIX 9: Re-description of a genetically typed, single oocyst line of the turkey coccidium, *Eimeria adenoeides* Moore and Brown, 1951 .................... 257

APPENDIX 10: Sequence-based genotyping clarifies conflicting historical morphometric and biological data for five *Eimeria* species infecting turkeys.... 270

APPENDIX 11: Simultaneous identification and DNA barcoding of 6 *Eimeria* species infecting turkeys using PCR primers targeting the mitochondrial cytochrome c oxidase subunit-1 (mtCOI) locus ........................................ 282
DECLARATION OF WORK PERFORMED

This research was supported by grants to John R. Barta from the National Sciences and Engineering Research Council of Canada (NSERC) and the Ontario Ministry of Agriculture Food/Ministry of Rural Affairs (OMAF/RA).

Mosunmola E. Ogedengbe was the recipient of the OVC PhD Program Scholarship, the Novartis Graduate Scholarship in Parasitology, Korean-Canadian Dr. F. Schofield Memorial Scholarship, Dr. Wilson Henderson Memorial Scholarship, the Gallant Custom Laboratories Anniversary Scholarship and American Society of Parasitology Travel Award.

The studies described in this thesis are original and the experiments were performed by Mosunmola E. Ogedengbe with the following exceptions: 1) all Sanger sequencing reactions were completed by technical staff at the Laboratory Services Division of the Animal Health Laboratory, University of Guelph as a fee for service; and, 2) single-oocyst cloning and propagation of some avian coccidia upon which some of the research in this thesis depended was completed by Dr. S. El-Sherry and Ms J. Cobean.

Technical assistance (limited PCR amplification and cloning of PCR products) was provided by: Ms. J. Whale and Dr. M. Hafeez.

Additional parasite isolates or DNA templates used in this work were kindly provided by Drs. H.D. Chapman (University of Arkansas); B. Chobotar, (Department of Biology, Andrew’s University, MI); J.P. Dubey and B. Rosenthal, (USDA, Beltsville, MD); D. Martin, (IDEXX Laboratories, Markham, ON); Y. Qvarnstrom, (CDC Atlanta, GA); C. Su, (Department of Microbiology, Knoxville TN); and S. Weese, (Pathobiology, University of Guelph, ON).

Intellectual contributions of Drs. S. Weese, R. Gregory, T. Gariepy and J.D. Ogedengbe improved the studies described herein. All work was carried out under the supervision of Dr. J. R. Barta at the Department of Pathobiology, University of Guelph.
"We know that we, too, are collections of cells that work together, kept harmonized by chemical signals. If an organism can control those signals — an organism like a parasite — then it can control us. And therein lies the peculiar and precise horror of parasites."

Carl Zimmer, August 2000

1. INTRODUCTION

There are parasitic forms or parasitism exhibited by most of the more than 30 animal phyla with a few exceptions among the chordates and echinoderms (Brooks and McLennan, 1993; Zimmer, 2000). Parasites make up the majority of eukaryotic species on Earth (Zimmer, 2000; Besansky et al., 2003) and parasites enjoy greater ecological success than other species (Pérez-Ponce de León and Nadler, 2010). Biologists estimate there are at least 8 million undescribed species compared with only 1.7 million named species and, of those described species, relatively little is known concerning their biodiversity (Stoeckle and Hebert, 2008). Precisely how many parasites are out there? The exact number of extant parasites is unknown. What is evident is that parasites inhabit practically every other living thing; for example, humans, domestic animals and wildlife each play host to at least one apicomplexan parasite (Morrison, 2009). Many parasitic protists of medical and veterinary importance belong to the phylum Apicomplexa. The phylum Apicomplexa contains about 300 described genera containing an estimated 6,000 named species; several orders of magnitude more species probably exist than have been described to date (Levine 1988; Perkins et al., 2000; Morrison, 2009). For protists generally, about 0.1% of extant species have been described formally to date (Adl et al., 2007; Morrison, 2009). Apicomplexa, being parasitic protists, are well adapted and specialized for various predilection sites in all body regions, including skin, brain, mouth, liver, muscles, blood, intestine, or reproductive tracts of their hosts.

Phylogenetic analyses place Apicomplexa with the free living and parasitic dinoflagellates and ciliates as close relatives. Collectively they form the large and diverse clade called the Alveolata (see Cavalier-Smith, 1991; Patterson and Sogin, 1992). Of all of the proposed higher protistan taxonomic groups, the Alveolata appears to be the most consistently supported in all phylogenetic analyses. The Apicomplexa are
more distantly related to the Stramenopiles (the water molds, brown algae, golden algae and diatoms) that are also members of the Chromalveolata (see Keeling et al., 2005).

The Apicomplexa is a monophyletic taxon (a phylum whose descendants all originate from a common recent ancestor) (Barta et al., 1991) although there is debate as to whether or not the dinoflagellate-like *Perkinsus* spp. should be included in the phylum (Siddall et al., 1997; Morrison and Ellis, 1997; Ellis et al., 1998). Within the phylum Apicomplexa there are a number of distinct, widely accepted groupings of biologically or morphologically similar organisms: coccidia; cryptosporidia; piroplasms; haemosporinids; haemogregarines; and, gregarines (Perkins et al., 2000; Morrison, 2009). These widely recognized groups of Apicomplexa are defined by morphological and biological characters shared among members of each group, (Perkins et al., 2000; Adl et al., 2007; Morrison, 2008; 2009).

The assignment of any species into a taxonomic group has traditionally been based primarily on phenotypic characteristics (e.g. shape, color, behavior and biochemical parameters as well as their relationship to their attendant hosts or vectors, tissue tropisms; prepatent and patent periods). Not surprisingly, the current taxonomic placements or arrangements within the Apicomplexa have been subject to many changes (e.g. Levine, 1985; 1988). Morphological structures do not necessarily reveal phylogenetically useful homologous characteristics. The consequence is that genera erected using morphological structures alone could become a ‘grab bag’ for species with uncertain classification because of the poor knowledge of the evolutionary relationships among the group (Simpson and Rogers, 2004). The remarkable variability displayed by species as seen in the different life cycle stages within the same species (Siddall, 1995) makes determination of character homology extremely difficult. The failure of morphological or biological characters to reliably assemble species into larger taxonomic units has not been solved yet by incorporating molecular analyses; the lack of clarity from more recent molecular analyses likely results from the relatively limited availability of molecular data from the Apicomplexa. Sequence data are frequently obtained from insufficient numbers or diversity of representative taxa, limiting the utility of any resulting phylogenetic analyses (Goloboff et al., 2009; Morrison, 2009). Even
with molecular data, generation of reasonable phylogenetic alignments that reflect assumed sequence and positional homologies is challenging and incorrect assumptions of homology can be reflected in the resulting phylogenetic trees (Barta et al., 1991; Kuo et al., 2008). Selection of suitable sequence alignment methods greatly influences the reliability of any resulting phylogenetic hypotheses (Crowe 1988; Morrison and Ellis, 1997). In recent times, phylogenetic relationships within the Apicomplexa have been defined and redefined (Adl et al., 2005) as researchers gained better understanding based on an accumulation of more comprehensive species descriptions supplemented with additional informative data through advances in molecular methods (Adl et al., 2005; Goloboff et al., 2009).

In the past three decades, molecular methods and the development of effective DNA identification techniques for use on a broader scale for species identifications and phylogenetic studies within the Apicomplexa have been the focus of many researchers (e.g. Barta et al., 1991; Cavalier-Smith, 1993; Tenter et al., 2002; Morrison et al., 2004; Morrison, 2008). Apicomplexan parasites have three distinct functional genomes that are accessible as genetic targets; excluding gregarines, in which a mitochondrial genome has not been found in all species belonging to this group so far. These three genomes have not been found in Cryptosporidium species, the plastid or mitochondrial genomes may have been secondarily lost. These multiple distinct functional genomes are the nuclear, mitochondrial and apicoplast genomes. Collectively, these genomes provide researchers options for exploiting different molecular loci that might help to better define phylogenetic relationships among parasites within the phylum (Wilson and Williamson 1997).

In molecular phylogenetic studies to date, the universality and abundance of targets within nuclear DNA and, in particular, 18S rDNA has made this genetic target widely used for species characterization and phylogenetic analysis (Adam et al., 2000, Gubbels et al., 2000; Cossio-Bayugar et al., 2002). Over 15,424 complete or partial 18S sequences have been submitted to GenBank from apicomplexan parasites but many of these are from a small number of species. For example, almost 8.2% (1260) of the 18S sequences in GenBank from apicomplexan parasites are from the single human pathogen
Toxoplasma gondii and 10 apicomplexan species (6 of which are human pathogens) account for over 32% (>4,862) accessions (http://www.ncbi.nlm.nih.gov/nuccore accessed 2015-02-19). Within the rDNA, the variable regions of the 18S rRNA gene and those of the more quickly evolving genetic targets of the nuclear ribosomal internal transcribed spacers (ITS-1 and ITS-2) have been investigated and used in studies of more closely related parasites. Although other nuclear genes have been used for molecular identification and phylogenetic studies (e.g. HSP70, β-tubulin, etc.), these studies frequently only cover a narrow taxonomic group (e.g. piroplasms - Yabsley et al., 2006a).

In addition to the nuclear genome, most members of the phylum Apicomplexa possess a functional unpigmented, non-photosynthetic 35 kB apicoplast (relic plastid, see Cai et al., 2003; Singh et al., 2003). Several hypotheses posit that apicomplexan plastids originated from the green algae lineages through secondary endosymbiosis (one symbiont dwelling within another symbiont that itself harbours a plastid) (Kholer et al., 1997; McFadden et al., 1997) or from red algae because gene content of apicoplasts bear a resemblance to those of the red algae (Wilson, 2002). Researchers exploiting the 35kb apicoplast genome have used various apicoplast gene sequences or whole genome sequences to help define and identify species of Apicomplexa e.g. Plasmodium species (Gadner et al., 1994), to determine phylogenetic relationships among Eimeria spp. (Zhao et al., 2001), Toxoplasma gondii and other tissue coccidia (Waller et al, 1998; Striepen et al., 2000; Zhang et al., 2000), and to study the origin of plastids in Apicomplexa and dinoflagellates (Zhang et al., 2000; Fast et al., 2001).

Mitochondrial genomes have been found in all apicomplexan taxa studied to date with the exception of cryptosporidia and gregarines. Apicomplexa possess a small 6kb mitochondrial genome that appears to be subjected to partial loss of sequence and in some cases complete transfer of the organellar DNA to the nucleus. Evidence of secondary loss of a mitochondrion has been seen in Cryptosporidium species which respire anaerobically through a greatly reduced respiratory chain (Mogi and Kita, 2010). In gregarines such as Pyxinoidea balani (see Reger, 1967; Clopton, 2002), Selenidium vivax (see Leander, 2006) and Gregarina niphandrides (see Toso and Omoto, 2007), a
mitochondrion was identified morphologically using transmission electron microscopy (TEM). In apicomplexan protists in which the mitochondrial genome has been identified, protein coding genes for cytochrome c oxidase subunit I (COI), cytochrome c oxidase III (COIII) and cytochrome oxidase b (CytB) are present. The three mitochondrial protein coding genes have been exploited in phylogenetic and related studies. More universally, the COI gene was promoted by Hebert et al. (2003; 2004) as a useful ‘DNA barcode’ gene (see www.ibol.org) and COI has been shown to be useful for species delimitation with various parasites (Ferri et al., 2009) and numerous free-living organisms (Hebert et al., 2003).

To some extent, the evolutionary relationships and taxonomic placements of parasites in this phylum remain unclear and inconclusive (Morrison, 2008) despite diverse morphological and molecular data available for inferring phylogenetic histories. Biodiversity and biological studies can benefit from a stable taxonomy that is based on accurate morphological and phylogenetic information. In turn, such accurate morphological and phylogenetic data can be expected to produce monophyletic taxonomic groups (e.g. species, genera, families, etc) arranged into an evolutionarily relevant hierarchy (Morrison, 2008; 2009). This ideal necessitates re-evaluating morphological and molecular phylogenetic information aimed at clarifying the evolutionary associations among species within the Apicomplexa (e.g. Barta et al., 1991; 2001; Morrison and Ellis 1997; Tenter et al., 2002; Simpson and Rogers 2004; Morrison, 2008). Some genes encoded by apicomplexan mitochondria, especially cytochrome c oxidase subunits I (COI) and III (COIII), appear to contain phylogenetically informative sequence variation that is not complicated by recombination events typically found with nuclear-encoded genes (Elson et al., 2001). The utility of these genetic loci for species identification and phylogenetic inference has been demonstrated for some haemosporinids, piroplasms and coccidia (Perkins and Schall, 2002; Brayton et al., 2007; Cunha et al., 2009; Ogedengbe et al., 2011).
2. LITERATURE REVIEW

2.1. BIOLOGY OF THE PHYLUM APICOMPLEXA

Members of the phylum Apicomplexa are unicellular, mostly intracellular parasites, some of which are causative agents for important, and sometimes devastating, diseases of humans and animals. The group is united by the possession of the apical complex (after which it is named). The apical complex comprises ultrastructurally visible organelles (i.e. rhoptries, micronemes and conoid with its associated apical rings and subpellicular microtubules) found in the motile stages of all members of the Apicomplexa. The secretory organelles (rhoptries and micronemes) are used in adhesion to host cell membranes and motility (micronemes) and during the process of host cell invasion (micronemes and rhoptries). The possession of this highly conserved suite of synapomorphic organelles collectively called the ‘apical complex’ strongly supports a shared evolutionary origin. Colpodellids, which are considered to be sister taxa to the Apicomplexa, have apparently homologous apical structures although these structures have not been organized as cohesively as within the Apicomplexa (Leander, 2008; Myl’nikova and Myl’nikov, 2009). This trait is shared with the earliest branching ‘true’ apicomplexan parasites Perkinsus species. The apical structure conoids are incomplete in Perkinsus species (Villalba et al., 2004).

Locomotion by members of the Apicomplexa is generally by gliding (e.g. Baum et al., 2006) or body flexion (e.g. Leander, 2008) with no change in form or apparent contractions of the body. Some life cycle stages (microgametes = male gametes) move using flagella, cilia are absent. Feeding is through micropores or directly through the plasmalemma (cell membrane) of the parasite. The feeding through the plasmalemma is sometimes enhanced through a modification of the interface between the parasites and their host cells into a structure termed a feeding organelle in cryptosporidia and gregarines (Valigurova and Koudela, 2005; Barta and Thompson, 2006). As far as is known, only vegetative or intracellular stages (e.g. tachyzoites, bradyzoites, merozoites, meronts, gamonts, gametes and trophozoites) feed via the micropores.

Most apicomplexan parasites are intracellular in host tissues. Their dependence on host cells for essential metabolites that they cannot produce for themselves ensure that
these parasites are difficult to cultivate in vitro and the vast majority have not been cultivated (Perkins et al., 2000). However, there are some exceptions when in vitro cultivation has been possible. Such instances include the use of snail tissue cultures for cultivating Klossia helicina (Moltmann, 1981) or the use of chicken, rabbit or cattle cells for cultivating some *Eimeria* and *Isospora* species (Doran, 1973). In other instances tissue coccidia such as *Toxoplasma* and *Sarcocystis* species were cultivated in both primary and established cell lines. Haemosporinids and piroplasms (e.g. *Plasmodium, Babesia* and *Theileria* species) were cultivated using host erythrocytes (Trager and Jensen, 1976; Brown, 1983). Except for reports of axenic culture of *Cryptosporidium* species (e.g. Hijjawa, 2010; Karanis and Aldeyarbi, 2011), in vitro culture of apicomplexan parasites is dependent on host cells. Some apicomplexan species are monoxenous (require only one host species to complete their life cycle) and others are either facultatively or obligately heteroxenous (requiring more than one species for their life cycle). Some parasites are highly host specific such as some coccidia (e.g. *Eimeria* spp. or *Sarcocystis* spp.) whereas others are more generalist in nature and are capable of infecting a wide range of hosts (e.g. *Toxoplasma gondii* or *Cryptosporidium parvum*). Barta (1989) suggested heteroxenous apicomplexan parasites probably evolved independently from ancestral monoxenous parasites. In parasites that later evolved to heteroxeny (sometimes optionally), the direct descendants of the ancestral host remained their definitive host while some or all of the asexual portion of the life cycle is moved to a second (intermediate) host that plays an ecological role in newly heteroxenous life cycle. Apicomplexan life cycles can be relatively simple or complex but most have asexual and all have sexual phases (Figure 2.1). Sexual reproduction is by syngamy (i.e. the union of two gametes to form a zygote). Microgametes (male) and macrogametes (female) fuse to form the diploid zygotes. The diploid zygote then initiates meiosis immediately to produce haploid progeny. Apicomplexan protists remain haploid throughout the remainder of their life cycles whether they undergo asexual replication or not. In gregarines and adeleid coccidia, gametogenesis is extracellular and is preceded by syzygy. Gamonts undergo syzygy (conjunction of macro- and microgamonts in a side-to-side or head-to-tail orientation) before final gamete maturation and fertilization. Whether the gametes mature in association (syzygy) or not, the zygote produces an
oocyst that contains sporozoites (Levine 1988). Asexual reproduction, when present, can occur by binary fission, schizogony (i.e. multiple fission of the nucleus followed by segmentation of the cytoplasm giving rise to multiple progeny) or by endodyogeny/endopolyogeny (i.e. separate nuclear divisions give rise to two or more daughter cells internally developed within a parent cell, e.g. in *Toxoplasma* and *Frenkelia* species) (see Perkins et al., 2000). There are no ‘sex chromosomes’ in apicomplexan parasites; in parasites with asexual replication, a single haploid sporozoite can replicate asexually and then develop both male and female gametes that can fuse to form a zygote (i.e., the haploid parasite can self-fertilize).

![Figure 2.1. Typical life cycles of apicomplexan parasites: The circle illustrates a basic apicomplexan life cycle consisting of both sexual and asexual developments. Three stages in the life cycle include: sporogony (I-L), merogony (A-D) and gametogony or gamogony (E-H). Adopted from Barta, 2001. Coccidiosis. Encyclopedia of life sciences © 2001, John Wiley & Sons, Ltd. www.els.net](image-url)
Some apicomplexan species are important pathogens of vertebrates and invertebrates and are responsible for significant morbidity and mortality in both human hosts and agricultural animals. Pathogenicity of parasites in this protistan phylum is frequently related to unlimited multiplication of their merogonic stages (Perkins et al., 2000). The haemosporinids and piroplasms cause significant economic impact on humans as well as domestic and wild animals. *Plasmodium* species are the etiological agents of malaria. The disease is amongst the most important human infectious diseases, with more than one million deaths per year in humans have been reported (WHO and UNICEF, 2005; WHO, 2008a; Manguin et al., 2010). Related *Haemoproteus* and *Leucocytozoon* species infect birds and lizards. Some of these parasites can cause severe economic losses during the production and rearing of ducklings, turkeys and other animals through morbidities and mortalities. Enteric coccidia in the suborder Eimeriorina (such as *Eimeria, Isospora* and *Cyclospora* species) cause self-limiting intestinal coccidiosis in a wide variety of birds and mammals. In particular, coccidia in poultry can produce substantial morbidity or mortality (Levine, 1988). Coccidiosis is associated with considerable economic costs reaching several billion USD annually in commercial poultry production (Williams, 1999). Tissue coccidia such as *Sarcocystis* and *Neospora* species are important parasites of domestic animals and ruminants. The tissue coccidium *Toxoplasma gondii* is responsible for congenital toxoplasmosis in both immunologically naïve humans and domestic animals resulting in infected fetuses, stillbirths and abortions (Montoya and Liesenfeld, 2004). Acquired (post-natal) toxoplasmosis is a serious disease of immunocompromised humans and animals. Pathogenic apicomplexan parasites including *Cryptosporidium* species are incriminated in cryptosporidiosis in immunologically compromised hosts (Abrahamsen et al., 2004). Among the Piroplasmorida are *Babesia* spp. that cause babesiosis, an important blood disease in ruminants (i.e. Texas cattle fever) as well as in humans (Brayton et al., 2007). *Theileria* species are agents of tropical theileriosis and the East Coast fever in cattle (Pain et al., 2005).

The gregarines Grasse 1953 are divided classically into three orders: archigregarines; neogregarines; and, eugregarines (Levine et al., 1980). Archigregarines are considered ancestral gregarines inhabiting hosts within marine waters. They undergo
merogony, gamogony and sporogony. Neogregarines likewise reproduce asexually through multiple merogonic cycles in the host and, like the haemosporinids and piroplasms, neogregarines can build up to large numbers in host tissues (Leander, 2008). Like the archigregarines, the neogregarines reproduce by merogony, gamogony and sporogony. Neogregarines parasitize insects typically and their replicative ability makes them potential pathogens in those hosts. Some species of the neogregarines such as Mattesia species have been considered as biological control agents of insect pests (Brooks, 1988; Lords, 2003). In contrast to the two gregarine orders discussed previously, eugregarines only undergo gamogony and sporogony (Leander, 2008). The eugregarines can be further divided into septate gregarines (depending on whether the trophozoite stage is separated into a protomerite and deuteromerite by a transverse septum, e.g. Gregarina species parasitizing insects) or aseptate forms in which trophozoites are not divided by the transverse septum (e.g. Monocystis species parasitizing earthworms) (Clopton, 2000).

2.2. CLASSIFICATIONS OF SUPER GROUPS

Classifications of organisms within the super groups (see Figure 2.2) into apparently monophyletic groups are proposed on a regular basis (e.g. Levine et al., 1980; Adl et al., 2005; 2012). The domain Eukaryota is divided into five super-groups: Opisthokonta; Amoebozoa; Excavata; Archaeplastida; and, SAR (Adl et al., 2012). ‘SAR’ is an acronym for the Stramenopiles, Alveolata and Rhizaria (Burki et al., 2008; 2012). Alveolata and Stramenopiles form a monophyletic group with the super phylum Rhizaria (Adl et al., 2012). ‘SAR’ became the new super phylum that replaced the paraphyletic (group does not contain all descendants from a common ancestor) Chromalveolata that contained the Haptophyta, Cryptophyceae, Alveolata and Stramenopiles (Adl et al., 2005; Burki et al., 2007; Hampl et al., 2009). A new higher classification scheme was necessitated based on analyses of a range of phylogenetic data (i.e. expressed sequence tags (Burki et al., 2007) and multigene analyses (Reeb et al., 2009) as well as morphological analyses (Adl et al., 2012) that gave support for the monophyly of the new super phyla. The super phylum Alveolata Cavalier-Smith 1991 comprises the ciliates, dinoflagellates and Apicomplexa. The super phylum
Stramenopiles comprises the brown algae, diatoms, opalinids and many fungi. The ancestor of the taxa now classified as belonging to SAR likely possessed a plastid obtained through secondary endosymbiosis with a cyanobacterium. Although plastid was secondarily lost or reduced in some species, it may have been reacquired in others (Adl et al., 2005). Within the Alveolata, the dinoflagellates and ciliates are sister taxa to the phylum Apicomplexa based on shared structural features. All alveolate taxa possess a tri-membranous surface layer; the cortical alveoli, an inner parallel membrane complex of flattened subpellicular vesicles after which the superphylum was named. This structure lies directly beneath the plasmalemma (outermost membrane) of the protist and provides additional structural support and rigidity to the cell. Alveolata also possess tubular cristae within their mitochondria that are characteristic to this superphylum (Cavalier-Smith, 1991; Leander and Keeling, 2004). Cavalier-Smith, (1993) postulated that the cortical alveoli appeared once in evolution and thus the alveoli located below the plasmalemma became a decisive defining feature for this super-phylum. The presence of the ultrastructural synapomorphies of open-sided conoids, alveoli and micropores (Siddall et al., 1997) and relationships inferred from molecular sequence data gave support to the superphylum as being monophyletic (Cavalier-Smith, 1991; 1993; Gajadhar et al., 1991; Kuvardina et al., 2002; Morrissette and Sibley, 2002; Harper et al., 2005). Dinoflagellates and apicomplexan parasites together make up the Myzozoa within the Alveolata. Members of the Myzozoa are united by their use of myzocytosis for feeding, possession of a haploid nucleus (except immediately following syngamy), presence of flattened cortical alveoli, rhoptries and micronemes (Fensome et al., 1999; Cavalier-Smith and Chao, 2004). The lack of a complete apical complex (Figure 2.3) was used to differentiate the dinoflagellates from the phylum Apicomplexa (Adl et al., 2012). Cavalier-Smith (1993) suggested the nonphotosynthetic phylum Apicomplexa must have had a photosynthetic distant ancestor because of the presence of a circular 35kb chloroplast genome found within the apicoplast organelle. Members of the phylum Apicomplexa have therefore evolved and adapted the use of the apical complex for their endoparasitic life.
Figure 2.2. A view of eukaryote phylogeny reflecting the five super-groups classifications namely: Opisthokonta, Amoebozoa, Excavata, Archaeplastida, and SAR, (acronym for three phylogenetic related groups, the Stramenopiles, Alveolata, and Rhizaria); from Adl et al., 2012 (J. Euk. Microbiol. 59:429-514, DOI: 10.1111/j.1550-7408.2012.00644.x), with permission.

Figure 2.3. Apical complex organelles. Diagrammatic representation of a typical apicomplexan zoite, showing the unique apical complex organelles found in motile stages of parasites in the phylum Apicomplexa: rhoptries, micronemes, conoids and apical polar ring. Image from Dr. J. Slapeta (see http://tolweb.org) modified under a Creative Commons Attribution-NonCommercial 3.0 unported license (CC BY-NC 3.0).
The current taxonomic structure of the phylum Apicomplexa Levine 1970 was proposed based on consensus in evidence presented from phylogenetic relationships deduced from several genetic markers. These genetic markers included the large and small ribosomal RNA subunits, transfer RNA, glyceraldehyde-3-phosphate dehydrogenase and superoxide dismutase and triosephosphate isomerase. Phylogenetic analyses were also complemented with relationships proposed based on morphological and ultrastructural characters (Van De Peer et al, 1993). Most motile forms of parasites in the phylum Apicomplexa Levine 1970 possess a specialized apical complex (from which the phylum got its name) consisting of the following subcellular organelles: polar rings; rhoptries; micronemes; conoids; and, sub-pellicular microtubules. Male gametes of some apicomplexan parasites move using either free or intracellular flagella (Levine 1988). The phylum Apicomplexa is divided taxonomically into two groups, Aconoidasida Mehlhorn Peters & Haberkorn 1980 and Conoidasida Levine 1988,
distinguished based on the presence or absence of a complete conoid within the apical complex. Members of the Aconoidasida typically lack conoids in motile stages with exception of some haemosporinids that express complete conoids in the ookinete stage (motile zygotes). Aconoidasida include the Haemospororida Danilewski 1885 (important parasites in the genera *Plasmodium, Haemoproteus, Leucocytozoon* and *Hepatocystis*) and the Piroplasmorida Wenyon 1926 (including parasites in the genera *Babesia, Theileria* and *Cytauxzoon*). Parasites within the Conoidasida, on the other hand, possess a complete apical complex including a completely developed conoid in all or most asexual motile stages. As illustrated in Figure 2.4, the Conoidasida is likely not monophyletic (Mehlhorn et al., 1980; Cavalier-Smith, 1993; Adl et al., 2005; 2012). There are two widely recognized lineages within the Conoidasida. The first, Coccidiasina Leuckart 1879, encompasses most of the coccidia including important species in the genera *Eimeria, Lankesterella, Toxoplasma, Sarcocystis, Cyclospora, Cystoisospora* or *Hepatozoon*. The Gregarinasina Dufour 1828 are parasites of invertebrates including species in the genera *Monocystis* and *Gregarina*. Genetic evidence (Carreno et al., 1999) and an apparent shared loss of both plastid and mitochondrial genomes (Slapeta, 2011) suggest that *Cryptosporidium* spp. (coccidia-like parasites that infect vertebrates) likely share a common ancestor with the gregarines (see Figure 2.4; Table 2.1). Species within the Coccidiasina produce gametes that fuse to form non-motile zygotes (oocysts) within which infective sporozoites develop (sometimes within sporocysts). Among the Conoidasida, gregarines have evolved distinct diverse surface structural patterns and are parasites of annelids, arthropods and molluscs. Gregarines inhabit extra-intestinal cavities, coeloms and reproductive vesicles of their hosts (Rueckert and Leander, 2008). *Selenidium* species (archigregarines) have retained certain ancestral qualities of exclusive marine habitation and possession of trophozoites that bear striking resemblance to infective sporozoites. Monocystids (e.g. *Lecudina* or *Monocystis* spp.) and urosporidians (e.g. *Pterospora*, aseptate eugregarines), possess large trophozoites that differs considerably from the sporozoite by loss of the apical complex. This loss is replaced by a mucron or epimerite. *Ophryocystis* and *Mattesia* spp. (neogregarines) have reduced trophozoite stages
Table 2.1. Taxonomic classification of the phylum Apicomplexa with some exemplar genera and species (after Levine, 1988).

PHYLUM APICOMPLEXA Levine, 1970

Class Perkinsasida Levine, 1978
Genus *Perkinsus* Levine, 1978

Class Conoidasida Levine, 1988
Subclass Gregarininasina Dufour, 1828
Order Archigregarinorida Grassé, 1953
  *Selenidioides* Levine, 1971; *Exoschizon* Hukui, 1939
Order Eugregarinorida Léger, 1900
  Suborder Blastogregarinorina Chatton and Villeneuve, 1936
  *Monocystis* von Stein, 1848; *Lankesteria* Mingazzini, 1891
  Suborder Septatorina Lankester, 1885
  *Gregarina* Dufour, 1828; *Stylocephalus* Ellis, 1912
Order Neogregarinorida Grassé, 1953
  *Ophryocystis* Schneider, 1883; *Schizocystis* Léger, 1900
Subclass Coccidiasina Leuckart, 1879
Order Protoeugregarinorida Kheisin, 1956
  *Grellia* Levine, 1973; *Coelotropha* Henneré, 1963
Order Eucoccidioida Léger and Duboscq, 1910
Suborder Adeleorina Léger, 1911
  Family Haemogregarinidae Léger, 1911
  *Haemogregarina* Danilewsky, 1885; *Hepatocystis* Miller, 1908
Suborder Eimeriorina Léger, 1911
  Family Cryptosporidiidae Danilewsky, 1885
  *Cryptosporidium* Tyzzer, 1907
Family Eimeridae Minchin, 1903
  *Eimeria* Schneider, 1875; *Cyclospora* Schneider, 1881; *Caryospora* Léger, 1904;
  *Isospora* Schneider, 1881
Family Atoxoplasmatidae Levine, 1982
  *Atoxoplasma* Garnham, 1950
Family Lankesterellidae Nöller, 1920
  *Lankesterella* Labbé, 1899
Family Sarcocystidae Poche, 1913
  Subfamily Sarcocystinae Poche, 1913
  *Sarcocystis* Lankester, 1882; *Frenkelia* Biocca, 1968
  Subfamily Toxoplasmatinae Biocca, 1957
  *Toxoplasma* Nicolle and Manceaux, 1909; *Besnoitia* Henry, 1913

Class Aconoidasida Mehlhorn, Peters, and Haberkorn, 1980
Order Haemosporina Danilewsky, 1885
Family Plasmodiidae Mesnil, 1903
  *Plasmodium* Marchiafava and Celli, 1885; *Polychromophilus* Dionisi, 1899;
  *Hepatocystis* Levaditi and Schoen, 1932 emend. Garnham, 1951;
  *Haemoproteus* Kruse, 1890; *Leucocytozoon* Sambon, 1908
Order Piroplasmodorida Wenyon, 1926
Family Babesiidae Poche, 1913
  *Babesia* Starcovici, 1893
Family Theileriidae du Toit, 1918
  *Theileria* Bettencourt, França, and Borges, 1907
(Rueckert and Leander, 2008). Other species of gregarines include *Lankesteria*, *Levinea*, *Menospora*, *Nematocystis*, *Nematopsis*, *Steinina*, and *Trichorhynchus* (Adl et al., 2005).

2.3. IDENTIFICATION METHODS AND CHALLENGES

2.3.1. MORPHOLOGICAL IDENTIFICATION

Classification of organisms has depended on the binomial Linnaean method. Systematic biologists and taxonomists dealing with ‘animals’ (Animalia or Metazoa) generally follow specific taxonomic guidelines (the ‘code’) overseen by the International Commission on Zoological Nomenclature (ICZN, 1999) that guide the formal naming of genera and species and their organization into family level groupings (Tautz et al., 2003; Frezal and Leblois, 2008). The formalized ICZN rules have enabled zoologists to differentiate and formally name an estimated 1.7 million species, accounting for about 10% of the estimated 10-15 million total species on earth. Hebert et al. (2003a) estimated that as many as 15,000 taxonomists would be required to routinely and continually identify organisms if our dependence on morphological diagnosis is to be sustained because not more than 0.01% of the estimated diversity of extant life on earth can be identified readily by a single taxonomist. The dwindling number of high level specialists cannot cover all taxa (Frezal and Leblois, 2008); this insufficient expertise and knowledge frequently leads to misidentification of specimens. Added to this is the difficulty of accessibility of experts’ information when they retire (Tautz et al., 2003) and dwindling number of newly trained taxonomists available to replace those approaching the end of their careers.

Morphological and morphometric characters used to define members of the phylum Apicomplexa have been obtained through the use of optical and electron microscopy with measurements following methods described by Levine, (1984; 1988) Duszynski and Wilber, (1997); Olson et al., (1998) and Martisen et al., (2006). Morphological characters utilized in classification of apicomplexan protists generally depended on the shape, size and color of various life history stages such as oocysts, sporocysts, sporozoites, gametocytes and their intracellular position, and subcellular organelles (Paperna and Landsberg, 1989b; Perkins et al., 2000). These characters are usually supplemented with life-cycle details including all possible hosts that can be
infected by a particular parasite (Valkiunas, 2005; Martisen et al., 2006). Other details may include predilection sites (i.e. tissue tropisms), prepatent and patent periods, sporulation times, gross or histopathological tissue changes in infected hosts, and even biogeographical data to identify parasites (Levine 1988; Paperna and Landsberg 1989b; Besansky et al., 2003). In some instances, characters with questionable utility, such as the species of host parasitized, may be included (Martisen et al., 2006). Morphological identification is not only tenuous but time-consuming; experimental infections may be required to observe detailed life cycle patterns before an organism can be correctly assigned to a particular species or even a higher taxonomic group. Given the limited taxonomic expertise available globally, it is not too surprising that there is a demonstrated bias towards identifying or characterizing groups of apicomplexan parasites that cause significant disease in companion animals, food production animals and humans. Thus, a majority of these parasites in infrequently examined or ‘unimportant’ vertebrate hosts remain undiscovered or, even if known, not fully described formally (Besansky et al., 2003; Tautz et al., 2003; Stoeckel and Hebert, 2008). This is an even greater issue with regard to apicomplexan parasites infecting invertebrates where our knowledge of their parasitic fauna is woefully inadequate (Levine 1988). The challenge to correctly identify species is compounded by the many morphologically dissimilar life history stages existing within a single apicomplexan parasite. These diverse forms are a challenge to unambiguous identification because phenotypic plasticity and limited distinguishing morphological characteristics combine to bedevil assignment to the correct species (or even genus). For example, haemogregarine gamonts (the accessible stage in the life cycle of these parasites because they can be observed in a simple blood smear) cannot be distinguished morphologically; these intraerythrocytic gamonts can belong to one of 5 or 6 haemogregarine genera, each with vastly different life cycles. The same is true with other parasites such as the asexual stages of many tissue coccidia (e.g. Toxoplasma, Neospora, or Hammondia) for which morphological differentiation is simply impossible (Besansky et al., 2003). Morphological identification methods are also limited because they may overlook the existence of cryptic taxa; such species have no qualitative morphological and/or
morphometric data available that permit their detection (Pérez-Ponce de León and Nadler, 2010).

2.3.2. MOLECULAR IDENTIFICATION, DIFFERENTIATION & PHYLOGENETICS

As an adjunct to the use of morphology-based (i.e. “morphospecies” defined by reference to a morphological type) taxonomic and systematic studies, molecular targets have been identified as alternative sources of data that can facilitate taxonomic studies. Tautz et al. (2003), Adl et al. (2005) and Barta and Thompson (2006) proposed the use of morphological details in conjunction with molecular data for species identification and classification. In order to validate the divisions and organizations of species within the Apicomplexa into taxonomic units, researchers have amplified and sequenced a range of available molecular targets. These genetic targets have been analyzed to address the unresolved phylogenetic relationships that arose from either the lack of data or the use of conflicting data in assigning taxonomic groupings. Apicomplexa have a wide variety of genetic targets to exploit because these parasites typically possess three independently evolving functional genomes: nuclear (8-64Mb with chromosomes); mitochondrial (~6kb genome copy size in various physical forms, (e.g. linear, circular, concatenated, flip-flop inverted); and, apicoplast (~35kb circular genome of a remnant, non-photosynthetic, plastid-like organelle). Unfortunately, not all apicomplexan parasites possess all three genomes. Both gregarines and Cryptosporidium species apparently lack the plastid genome (Zhu et al., 2000; Obornik et al., 2002; Abrahamsen, 2004; Barta and Thompson, 2006; Toso et al., 2007). Additionally, Cryptosporidium species have lost their mitochondrial genome but at least some of the respiratory functions of the mitochondrion have been replaced by degenerate mitochondria containing no DNA (i.e. mitosomes or hydrogenosomes) (Abrahamsen et al., 2004; Hjort et al., 2010). The presence of extra-nuclear DNA encoded by mitochondria or apicoplasts or both provides additional data that can be exploited to infer relationships among apicomplexan parasites. The nuclear genome is subject to recombination events because of the sexual reproduction of apicomplexan parasites during their life cycles. Both the mitochondrial and apicoplast genomes are inherited uniparentally through the female gamete. Recombination events and paralogous genes are likely to be uncommon
in these maternally inherited genomes (Hebert et al., 2003a); therefore, sequences derived from these extrachromosomal DNA sources might be more reliable and accessible indicators of the evolutionary history of these parasites.

Sequences obtained from several nuclear genes, particularly the ribosomal small subunit RNA genes (18S rDNA) but also large subunit (LSU), internal transcribed spacer 1 (ITS-1) and 2 (ITS-2) rDNA, have found extensive utility in species delimitation and phylogenetics with a wide range of apicomplexan parasites. Additional molecular methods that do not target a specific gene such as RAPD-based sequence characterised amplified regions (SCAR) markers are found to be extremely useful for differentiating closely related apicomplexan parasites (e.g. Fernandez et al., 2003b; Ogedengbe et al., 2009). The small subunit ribosomal RNA gene (18S rDNA) sequences are well suited for inferring relationships among apicomplexan genera and the more quickly diverging ribosomal ITS regions were found to be useful in differentiating closely related species and, in some instances, strains (see Table 2.2).

Table 2.2. Molecular targets used with various groups of Apicomplexa

<table>
<thead>
<tr>
<th>GENOME</th>
<th>LOCI</th>
<th>TAXONOMIC GROUP</th>
<th>SPECIES</th>
<th>REFERENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleus</td>
<td>18S rDNA</td>
<td>Sarcocystidae</td>
<td>Sarcocystis sp.</td>
<td>Mathew et al., 2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Toxoplasma sp.</td>
<td>Carreno et al., 1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Neospora sp.</td>
<td>Ellis et al., 1994, 1995, 1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adeleorina</td>
<td>Hepatozoon canis</td>
<td>Monteneiro et al., 2007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>H. americanum</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>H. catesbianae</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RAPD-SCAR markers</td>
<td>Gregarines</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plastid genes</td>
<td>Cryptosporidium</td>
<td>Cryptosporidium sp.</td>
<td>Leander et al., 2003</td>
</tr>
<tr>
<td></td>
<td>ITS</td>
<td>Eimeriorina</td>
<td>Eimeria sp.</td>
<td>Clopton, 2009</td>
</tr>
<tr>
<td></td>
<td>TRAP-C2, β-tubulin, 60 kDa glycoprotein (gp60)</td>
<td>Haemoproteida</td>
<td>Plasmodium sp.</td>
<td>Xiao, 2010</td>
</tr>
<tr>
<td></td>
<td>Mitochondrion mt COI</td>
<td>Eimeriorina</td>
<td>Eimeria spp.</td>
<td>Fernandez et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hepatozoon canis</td>
<td>Ogedengbe et al., 2009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>H. americanum</td>
<td>Jenkins et al., 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cantacessi et al., 2008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Boulianne et al., 2007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sulaiman et al., 1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Xiao, 2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Schwarz et al., 2009</td>
</tr>
</tbody>
</table>
However, there are limitations in the use of sequence data derived from the nuclear ribosomal small subunit (18S rDNA) targets. Marked limitations are imposed by the apparently incongruous observations of extensive sequence variability (in limited regions of the rDNA) making it difficult to infer positional homology in aligned sequences and the overall slow rate of evolution of the gene. This limitation reduces the ability of 18S rDNA sequences to differentiate closely related species (Barta et al., 1997; Li et al., 1997; Bhoora et al., 2009; Heger et al., 2011). Initially it was thought that the internal transcribed spacer (ITS) regions could address the problem of sequence variability and the difficulty of attaining positional homology preponderant in 18S rDNA sequences. These nuclear ITS regions are multiple copy gene targets that are not under strong discriminating pressure; for this reason they may permit identification of species and strains. However, ITS-1 and ITS-2 sequences proved to be too variable to provide phylogenetic information on the relationships among species, often even within the same genus (Lew et al., 2003). In addition, apicomplexan protists possess as many as three paralogous rDNA arrays within the nuclear genome (some of which are differentially expressed in different life cycle stages, e.g. McCutchan et al., 1988). Genetic distances between paralogs within a single species are frequently greater than the difference between homologs of two different species of the same genus. Clearly, inferring phylogenetic or taxonomic relationships using such confused sequence data becomes problematic. Finally, the 18S rDNA, ITS-1 and ITS-2 regions all utilize near universal PCR amplification primers that can amplify host DNA as well as eukaryotic contaminants within samples. As such, purified parasite material is preferred for successful PCR amplification of these genetic targets from the parasites to ultimately obtain reliable sequence data, something that can be challenging. Even with purified parasite DNA, PCR amplification can generate a considerable number of artifactual products (amplicon chimeras) that result from the fusion of sequences from the divergent rDNA arrays found in even a single apicomplexan nuclear genome (e.g. Miska et al. 2010; Cook et al. 2010; El-Sherry et al., 2013). Despite the attractiveness of the apparent universal applicability of rDNA and related ITS regions as data sources for DNA diagnostics to elucidate important pathogens, the aforementioned limitations have constrained the use of nuclear ribosomal genetic targets for many molecular
investigations. The features of the mitochondrial genome of apicomplexan parasites have drawn interest for the use of mitochondrial genes, particularly the cytochrome c oxidase subunit I (COI) in DNA barcoding of Apicomplexa (species level identification and delimitation). A growing database of sequences from apicomplexan mitochondrial genomes (Ondrejicka et al., 2014) promises well-annotated sequence data for inferring taxonomic and phylogenetic relationships in this taxonomically diverse phylum (Ogedengbe et al., 2011).

2.4. PHYLOGENETIC RELATIONSHIPS IN THE APICOMPLEXA

The phylum Apicomplexa is a significant and large taxonomic group with an estimated 6,000 members. Despite their importance as agents of disease, little is known regarding the evolution, biodiversity and interrelationships among the morphologically and biologically diverse taxa within this phylum. As Morrison (2009) explained, this frustrating lack of knowledge exists for a number of reasons: 1) difficulty in finding these parasites; 2) limited or incomplete understanding of their biology (particularly complete life history patterns); 3) ultrastructural features with uncertain homology; and, 4) uneven sampling of taxa within the phylum (a feast or famine availability of information on specific taxa determined largely by whether or not the parasite is of obvious economic or medical importance to humans). The resultant limited taxon sampling, lack of focused or directed collection of data for parasites in the phylum and difficulty of obtaining funding for research on ‘unimportant’ species have hampered efforts to better understand the biological diversity within the Apicomplexa. These obstacles have affected the ability to obtain a well resolved evolutionary history for apicomplexan parasites (e.g. Barta 1989; Morrison, 2009). Addressing the issues outlined above (see Morrison, 2009) may generate data that could be useful for identifying apicomplexan species reliably and such data may also be useful for inferring evolutionary relationships among apicomplexan parasites through rigorous phylogenetic studies. It is likely that additional data may support the existence of previously unknown (cryptic) species. The lack of a well supported evolutionary hypothesis for members of the phylum Apicomplexa has resulted in necessary revisions to long established taxonomic schemes. For example, Morrison, (2009) discussed how
Cryptosporidium was initially classified as a coccidian parasite (Perkins et al., 2000) but was found, in the light of additional phylogenetically informative data, to have closer taxonomic affinities with the gregarines (Carreno et al., 1999; Barta and Thompson, 2006; Leander, 2007).

The monophyly of the phylum Apicomplexa is reflected in their possession of a synapomorphic (shared, derived) ultrastructural feature - the apical complex. Although monophyly of the phylum can be supported using morphological features, inferring the evolutionary relationships among the various groups of apicomplexan taxa based on morphology becomes highly challenging because of the paucity of features for which homology can be reasonably assumed (Barta, 1989).

Molecular phylogeny deals with identifying and understanding evolutionary histories and relationships among extant and/or extinct species (descendants) in relation to a common descent (ancestor) based on changes in nucleic acid sequences in all or part of their genomes. Changes observed in sequences of orthologous genes (i.e. genes in different species that evolved from a common ancestral gene by speciation) as well as paralogous genes (i.e. genes related by duplication within a genome) are indicative of genetic divergence and, thus, are used to infer evolutionary history of the genes and the organisms over the course of time (Hoef-Emden, 2004). Phylogenetic trees are constructed from rigorously aligned homologous sequences obtained through sequencing of macromolecules (Hall, 2001; Gregory, 2008; Morrison, 2008, 2009). Hypothesized phylogenetic relationships are indicated by the branching patterns on the resulting tree(s). Subsequently, phylogenetic relationships based on these genes (whether orthologs or paralogs) are the basis for the current phylogenetic hypotheses for apicomplexan taxa. The Apicomplexa is a monophyletic group if Perkinsus species (parasites of mollusks and microeukaryotes and sister group to the dinoflagellates) are excluded from the phylum (Goggin and Barker, 1993; Ellis et al., 1998; Saldarriaga et al., 2003; Leander and Keeling, 2004). Perkinsus is currently the only genus in the ‘Perkinsea’ (an unranked taxonomic group within the Alveolata). Perkinsus species are differentiated morphologically from other apicomplexan species by the possession of an incompletely developed conoid within their apical complex; members of the
Apicomplexa possess symmetrical and complete conoids in all stages that possess this structure. Many authors seem to favor this reclassification; more recent molecular data suggest that *Perkinsus* spp. have stronger phylogenetic affinities with the phylum Dinoflagellata than with Apicomplexa (Levine 1988; Barta et al. 1991; Escalante and Ayala 1995; Siddall et al., 1995; 1997; Reece et al., 1997). Some researchers include the predatory zooflagellates of the genus *Colpodella* in the phylum Apicomplexa, although members of this genus are not parasitic. *Colpodella* species possess ultrastructural features that appear to be homologous to the apical complex found in members of the phylum Apicomplexa (see Cavalier-Smith, 1993); however, *Colpodella* species also possess flagella in their free-living forms (e.g. Simpson and Patterson, 1996). Others such as Kuvardina et al. (2002) consider *Colpodella* species as the nearest identified sister lineage to the ‘true’ apicomplexan parasites.

Barta et al. (1991) reported that the coccidia (Conoidasida; Coccidia) are monophyletic if members of the genus *Cryptosporidium* are excluded; this hypothesis was supported in a number of analyses based on the nuclear ribosomal small subunit gene (e.g. Carreno et al., 1999; Morrison, 2009). Nuclear 18S rDNA sequence-based analyses placed *Cryptosporidium* species as a sister group to all other Apicomplexa (Escalante and Ayala, 1995). Carreno et al. (1999), Barta and Thompson (2006), Leander (2008) and Morrison (2009) do not share this conclusion; *Cryptosporidium* spp. were found to group with the gregarines consistently if sequences from ‘minor’ apicomplexan parasites were included. This conclusion is at odds with classical taxonomic schemes of Levine (1988) and Corliss (1994).

Once the cryptosporidia were no longer considered ‘coccidia’, the classical taxonomic scheme for the subclass coccidia (Coccidia; Eucoccidiorida containing both adeleorinid and eimeriorinid coccidia in separate subclasses) is in good agreement with the molecular data. Coccidia in the families Eimeriidae and Sarcocystidae form well-supported monophyletic clades that are sister groups (Levine 1988; Barta et al., 1991; Ellis et al., 1994; 1995). Classical taxonomic schemes also agree with the molecular data at the subfamily level within the Sarcocystidae. The Sarcocystinae and
Toxoplasmatinae have been demonstrated to be monophyletic based on nuclear 18S rDNA sequence data (e.g. Morrison (2009); Figure 2.5).

Haemogregarines (Coccidia; Eucoccidiorida; Adeleorina) are a speciose group of intracellular parasites in the circulating erythrocytes of vertebrates. Over 300 described species (Barta and Desser, 1989; Siddall, 1995) are assembled into eight genera: Haemogregarina Danilewsky 1885; Karyolosus Labbé 1894; Dactylosoma Labbé 1894; Hemolivia Petit, Landau, Baccam and Lainson 1990; Hepatozoon Miller, 1908; Babesiosoma Jackowski and Nigrelli 1956; Cyrilia Lainson 1981 and Desseria Siddall 1995. All haemogregarine species with a described life cycle are transmitted by haematophagous invertebrate definitive hosts. However, details of the life histories of the vast majority of species belonging to these genera are not completely known. Consequently, many incorrectly defined species, sometimes with unrelated biological characters are grouped together. In terms of phylogenetic relationships, not much is known about haemogregarine taxa. The associations proposed are based mainly on morphological characters (Barta, 1989; Siddall, 1995; Mathew et al., 2000) with few molecular data to assist in generating a phylogenetic hypothesis for taxa within the group (Lang-Unnasch et al., 1998; Carreno et al., 1999; Mathew et al., 2000; Perkins and Keller, 2001). With the limited number of sequences available, the genus Hepatozoon appeared basal to piroplasms in a weakly supported (50% BP value) subclade that included haemosporinids, suggesting that the genus Hepatozoon is paraphyletic (Lang-Unnasch et al., 1998; Carreno et al., 1999). The recognition that Hepatozoon species were biologically variable and that monophyly of the genus was unlikely was concluded previously based on morphological and life cycle features (Smith, 1996). In limited work using 18S rDNA sequences, Adelina spp. formed the sister group to a number of Hepatozoon spp. Aggregata spp. formed the sister group to the Adelina/Hepatozoon clade (Kopecka et al., 2006). Dactylosoma and Babesiosoma species (family Dactylosomatidae) show biological similarities with the piroplasms in higher vertebrates with development of sporogonic stages in intestinal cells and merogonic stages in salivary cells of their definitive invertebrate hosts (Barta and Desser, 1989; Barta, 1991); nonetheless, members of the Dactylosomatidae have been confirmed to belong to the Adeleina based on 18S rDNA sequence analysis (Barta et al.,
Figure 2.5 represents currently understood evolutionary relationships and branching order among widely recognized groups of apicomplexan taxa deduced mainly from phylogenetic analysis of 18S rDNA sequences (Barta and Thompson, 2006).

Studies on evolutionary relationships among gregarines (Conoidasida; Gregarinia) inferred from small subunit ribosomal DNA sequences (Leander et al. 2006; Leander 2007; Rueckert and Leander, 2008) and morphological and biological characters suggest that phylogenetic relationship among species within the eugregarines and archigregarines is far from well resolved. Paraphyly exists among species classified into these groups e.g. sequences obtained for the archigregarine, *Selenidium terebellae*, was found to cluster within a weakly supported clade with those of the eugregarine, *Lecudina tuzetae* and two aseptate eugregarines, *Lankesteria chelyosoma* and *Lankesteria cystodytae* (Leander et al., 2003a; 2003b; Rueckert and Leander, 2008).

![Figure 2.5](image-url)
Monocystis species (aseptate eugregarines), form sister taxa with the neogregarine parasite, Ophryocystis elektroscirrha. Although species within the genus Lecudina are paraphyletic, they are closely related to the monophyletic clade of Lankesteria species. Lecudina branched basally to a monophyletic clade of urosporidian gregarines (e.g. Pterospora). The Gregarina and Leidyana species (septate eugregarines) studied thus far form monophyletic clades.

Our understanding of the phylogeny of gregarines is only rudimentary because of the limited number of taxa that have been studied. Nuclear small subunit rDNA sequence data from only about 25 of the 1650 gregarine species belonging to 250 named gregarine genera have been obtained (Clopton, 2000). Thus, molecular data are lacking from the vast majority of described species (Leader et al., 2003; Rueckert and Leander, 2008; Clopton, 2009). In general, phylogenetic data support a close and perhaps sister relationship of Cryptosporidium spp. to the gregarines (Leander et al., 2003; Barta and Thompson, 2006). At the very least, both the gregarines and cryptosporidia are early branching lineages within the phylum Apicomplexa. Regardless, the previous taxonomic placement of Cryptosporidium species within the subclass Coccidia is not supported and members of the genus should no longer be considered ‘coccidia’.

Morrison and Ellis (1997) proposed that the class Aconoidasida is monophyletic which was reflected in taxonomic schemes erected by Levine (1985) and Corliss (1994). However, Barta et al. (1991) excluded the order Hematozoea from the class Aconoidasida and instead suggested a sister group relationships to Cryptosporidium species. Limited taxon sampling in the latter study (in particular, the lack of any gregarine sequences in the analyses) makes drawing conclusions regarding the relationships among major apicomplexan lineages problematic. The order Piroplasmida, containing species in the genera Babesia (112 species) and Theileria (39 species) (see Peirce, 2000), is monophyletic based on 18S rDNA sequences and forms a sister group to the haemosporidians (Morrison, 2009). However, there is paraphyly within the genus Babesia (see Ellis et al., 1992; Mackenstedt et al., 1994). Using the ribosomal 18S rRNA gene sequences of six Babesia, three Theileria and one Cytauxzoon species, Allsopp et al. (1994) inferred that four Babesia spp. sensu stricto (e.g. B. bigemina, B. caballi, B.
bovis and B. canis) formed a monophyletic clade. However, a second monophyletic clade contained the theilerian protist, Cytotauxzon felis, two Babesia species (i.e. B. rodhaini and B. equi) and the Theileria spp (Cox, 1994). Paraphyly within the genus Babesia was reported similarly by Ellis et al. (1992), Mackenstedt et al. (1994) and Yabsley et al. (2006b).

2.5. DNA BARCODING AND MITOCHONDRIAL GENOME SEQUENCING

2.5.1. MITOCHONDRIAL ORIGIN AND EVOLUTION

“Life is the interplay between structure, energy and information” Wallace, (2007).

The mitochondrion functions in oxidative phosphorylation which is fundamental in all eukaryotic cells that respire aerobically. A process of energy production in the form of adenosine tri-phosphate (ATP) occurs by oxidizing calories from organic substrate through series of interrelated steps including proton gradient (∆P) production to final oxidation that results in ATP production (Gray et al 1999; Andersson et al., 2003). Other functions associated with mitochondria beyond their primary role of ATP production include production of hormones, generation of heat, regulation of reactive oxygen species, control of apoptosis and regulation of cell cycles (Burger et al., 2003; McBride et al., 2006). Mitochondria contain their own DNA-based genome that can encode for structural RNAs and various proteins, but the number and identities of the genes varies widely among species. Some protists have as many as 97 identifiable genes (e.g. Reclinomonas americana, a heterotrophic flagellated protist) and others can have as few as 5 (e.g. Plasmodium falciparum, an apicomplexan protist). Mitochondrion-encoded proteins such as cytochrome c oxidoreductase (CytB), cytochrome c oxidase subunits I and III play important roles in the respiratory pathway. Other respiratory enzymes that are encoded include: NADH: ubiquinone oxidoreductase encoded by nad; succinate: ubiquinol oxidoreductase encoded by sdh; and, ATP synthase (Gray et al 1998; 1999; Burger et al., 2003). In addition to enzymes, small subunit (SSU) and large subunit (LSU) ribosomal RNAs as well as tRNAs are encoded by some mitochondrial genomes. The mitochondrion in eukaryotes is thought to have an endosymbiotic origin (Gray, 1999). Mitochondria in eukaryotic cells are believed to be derived from a α-proteobacterial endosymbiotic ancestor. Phylogenies inferred from 16S
rDNA sequences of various bacteria and 16S rDNA sequences from eukaryote mitochondria (Yang et al., 1985 and Gray, 1999) suggested that the mitochondrion was derived from a $\alpha$-proteobacterium, likely within the Rickettsiales; many members of the Rickettsiales live in endosymbiotic relationships within eukaryotic cells. Using sequence data from protein-coding genes, Wallace (2007) supported the descent of the mitochondrion from a single protomitochondrial progenitor sharing ancestry with $\alpha$-proteobacteria. Thus, mitochondria in eukaryotic cells were derived from the $\alpha$-proteobacterium that developed an irreversible fusion with an anaerobic cell (Park and Larsson, 2011).

Although functionality is widely conserved, mitochondrial gene contents and arrangements, genome sizes, and gene expression vary dramatically among eukaryotes. Mitochondrial genome size ranges from the approximate $\sim$6 kB (kilobases) in *Plasmodium* to $\sim$367 kB in the flowering plant *Arabidopsis* (Marienfeld et al., 1999; Unseld et al., 1997). The largest mt genome known to date is $\sim$11,318 kB found in the flowering plant *Silene conica* (Sloan et al., 2012). All mitochondrial genomes have protein-coding genes that lack introns and use a genetic code distinct from eukaryotes (see Gray et al., 1998; 1999). The so-called “ancestral” (i.e. original form) mitochondrial genomes are typically circular and contain many genes, including the following: *nad; sdh; atp;* ribosomal protein genes (*rpl* and *rps*); large subunit (LSU or 23S), small subunit (SSU or 16S) or 5S rRNAs; and, tRNAs. Substantially “derived” (modified) mt genomes are reduced from the ancestral mitochondrial genome and lack many of the genes found in the ancestral forms. In addition, derived mt genomes exhibit considerable structural diversity. Their evolutionary development is marked by reduction in the number of protein coding genes, reduction or elimination of tRNAs, diversity in gene order and arrangements, rapid rate of genetic divergence and a considerable diminution in overall size (Gray et al., 1999; Lang et al., 1999).

**2.5.2. APICOMPLEXAN MITOCHONDRIAL GENOMES**

The apicomplexan mt genomes are among the extreme of these “derived” type mt genomes in eukaryotes (Gray et al., 1999; Feagin, 2000; Gray, 2012). The apicomplexan mitochondrial (mt) genome is the most highly reduced of the ‘derived’ mt
genomes (sensu Gray et al., 2004). The genome is also one of the smallest known at only 6-11 kilobases (Feagin, 1992; Feagin, 2000; Gray et al., 2004; Gray, 2012).

The physical form of apicomplexan mt genomes is variable (Figures 2.6-2.8). Apicomplexa mt genomes have been described as circular or tandemly repeated genomes in Plasmodium spp. (see Feagin et al., 1991; Wilson et al., 1991; Feagin, 1992; Preiser et al., 1996; Wilson and Williamson, 1997). In Plasmodium species, mitochondrial genomes occur as tandemly repeated copies, each about 6kb in length, arranged in a head to tail configuration (Vaidya and Arasu, 1987). This linear concatemeric form of the mt genome is predominant in most Plasmodium species (Feagin, 2000) and ensures that the mt genome map for these parasites is circular-mapping. In this form, the repeating mitochondrial sequences can then be cleaved at one or more locations with restriction endonucleases, producing fragments of varying sizes. Restriction mapping of the mt genomes of several Plasmodium species confirmed the presence of tandemly duplicated or circular genomes (e.g. Vaidya and Arasu, 1987; Joseph et al., 1989). Restriction fragments from Plasmodium species failed to hybridize to sequences from non apicomplexan organisms such Trypanosoma brucei or Saccharomyces cerevisiae (Vaidya and Arasu, 1987) or hybridized weakly to other apicomplexan parasites (i.e. Babesia, Theileria, Toxoplasma or Eimeria species) (e.g. Joseph et al., 1989).

Parasites in the piroplasmid genera Babesia and Theileria species possess 6.6 kb to 8.2 kb linear monomeric mt genomes with terminal inverted repeats (Kairo et al., 1994; Omori et al., 2007; Lau, 2009; Hikosaka et al., 2011).

The more distantly related eimeriid coccidia in the genus Eimeria (E. acervulina, E. brunetti, E. maxima, E. necatrix, E. tenella and E. praecox - etiological agents of coccidiosis in poultry) have genomes approximately 6kb in length (Lin et al., 2011). In these species, the mt genomes occur as linear concatemers (i.e. tandemly repeated) genomes (see Hikosaka et al., 2010; Lin et al., 2011).
Regardless of the structural or physical form of mt genomes in the phylum Apicomplexa, the actual genetic content of apicomplexan mitochondrial is remarkably conserved. Apicomplexan mt genome content usually consists of three protein-coding genes encoding COI, COIII and CytB, and highly fragmented large subunit (LSU) and small subunit (SSU) ribosomal RNA (rDNA) genes. The fragmented ribosomal RNA genes are found interspersed among the 3 protein-coding genes (Wilson et al., 1991; Hikosaka et al., 2010; Lin et al., 2011). Apicomplexan mt genomes completely lack 5S rRNA and tRNA genes (Gray et al., 1999) (see Figures 2.6 to 2.8). Although the two piroplasmid mitochondrial sequences had fewer (6) annotated LSU fragments compared to the 19 rDNA fragments identified in Plasmodium (Kairo et al., 1994; Feagin et al., 1997; Lau, 2009; Hikosaka et al., 2010), it is possible that the number of rDNA fragments was similar in both genomes but that the annotations differed. The unusual genome organization of some piroplasms (e.g. Theileria equi, see Hikosaka et al., 2010) includes some gene duplication, particularly towards the ends of their linear genomes. All other apicomplexan mt genomes examined to date lack any obvious gene duplications.
Figure 2.7. Diversity in the linear mitochondrial genomes of *Babesia gibsoni* (a), *Theileria orientalis* (b), *Theileria equi* (c) and *Theileria parva* (d). Variability in transcriptional directions of genes within the mt genomes of piroplasms was common. A flip-flop direction between COI and CytB, and inversions in all three protein coding genes in two *Babesia* species and three *Theileria* spp mt genome sequences was reported by Hikosaka et al. (2010).

Figure 2.8. Diagrammatic representation of a single mitochondrial genome ‘unit’ of *Eimeria* species. Although circular-mapping, mitochondrial genomes of *Eimeria* species are found as linear concatemers of several of the illustrated genome units. Concatenated mitochondrial genomes are found commonly within the Eimeriidae (Lin et al., 2011, *Gene* 408: 28–33)

The order or transcriptional directions of genes within mt genomes of different apicomplexan parasites can vary considerably (Figures 2.7(a-d), Feagin et al 1992; Hikosaka et al. 2010; 2012). Usually, such variability has been observed only between species or, more commonly, between major groups of parasites such as malarial parasites and coccidia. However, Hikosaka et al. (2012) reported an apparently stable flip-flop recombination phenomenon in some *Babesia* species that resulted in four distinct mt genome structures within a single parasite species, each with the same genes mapped uniquely.

What accounts for re-ordering of genes in mt genomes of apicomplexan parasites? One possible explanation could relate to the lack of an apparent ingenious DNA repair process in mitochondria. Unlike DNA in the nuclear genome, mitochondrial DNA is not protected by the presence of histone proteins (Remmen and Richardson, 2001; Prigione
et al., 2011). The mitochondrion is the major producer of free oxygen radicals (e.g. the reactive oxygen species, ROS) during oxidative phosphorylation (Cui et al., 2012a). ROS make the mtDNA susceptible to oxidative damage when reacting with mt macromolecules leading to damage to the sugar-phosphate backbone connecting the bases (Bogenhagen, 1999). The lack of histone may be permissive to the reorganization or resorting of fragmented mt DNA into novel new genome organizations. Such resorting and rearranging of mt genomes has apparently happened multiple times within the Apicomplexa (e.g. DeBarry and Kissinger, 2011). However, within most of the widely recognized groups in the phylum (e.g. the eimeriid coccidia or the haemospororids), the gene orders and directions are comparatively conserved; exceptionally, the piroplasms have a diverse range of genome organizations and lengths (DeBarry and Kissinger, 2011; Hikosaka et al, 2012).

In eukaryotes, there is compelling evidence that at least some of the nuclear DNA has been derived from mt DNA that was transposed and permanently incorporated into the nuclear genome (Gjerde, 2013a). Sequences that originate from the insertion of mtDNA into the nuclear genome are designated nuclear mtDNA or NUMTs (Lopez et al., 1994). NUMTS in Apicomplexa has been observed in some tissue coccidia especially in *Toxoplasma gondii*. Gjerde, (2013a) described more than 20 numts combinations from CytB and COI genes of *Toxoplasma gondii* flanked on both ends by non- coding sequences between them. Nuclear mtDNA can be identified by sequence similarity to the mitochondrial genome in the same organism.

Precisely why mt DNA are transposed to the nucleus is unclear. One suggestion is that mtDNA leakage into cytoplasm ends up in the nucleus. The simplification from ancestral to derived mt genomes (sensu Gray et al., 1998) is well demonstrated by the reduction in the complement of genes encoded by the mt genomes of apicomplexan parasites. Most apicomplexan parasites retain only three protein-coding genes (COI, COIII, CytB) and fragmented remnants of the 16S (SSU) and 23S (LSU) rDNA genes (Feagin 2012; Gray et al., 2004). The reduction in gene content may have gone even further in some of the coccidia. Even the greatly reduced mt genome of other Apicomplexa has not been detected in the tissue coccidium *Toxoplasma gondii*. However, sequences apparently homologous with two typical mt-encoded genes, COI
and CytB, have been identified within the nuclear genome (Ossorio et al., 1991). Cryptosporidium species may be the most extreme example of loss of mtDNA in the Apicomplexa. Based on molecular phylogenetic analyses (e.g. DeBarry and Kissinger, 2011; Barta and Thompson, 2006; see Figure 2.5). The cryptosporidia likely branched early from the ancestral apicomplexan lineage (which presumably had a mt genome, albeit reduced, because the chromalveolate ancestors all share mt genomes) and yet cryptosporidia show no evidence of retaining a mitochondrial genome (Mogi and Kita, 2010; Figure 2.4). There is no functional mtDNA in Cryptosporidium; limited respiratory functions of these parasites are fulfilled by a mitosome (or hydrogenosome) that utilizes a unique mitochondrial protein import pathway (Alcock et al., 2012). All mitosome-related proteins are encoded by the nuclear genome (Mogi and Kita, 2010).

In Eimeria species, the genome consists of three protein coding genes COI, COIII, and CytB, as well as a partial, fragmented large subunit (LSU) rRNA gene (12 fragments) and 7 fragments of the small subunit (SSU) rRNA gene. No genes encoding transfer RNAs are present. Putative translational direction for the protein genes was the same in all Eimeria species. Both Hikosaka et al. (2010) and Lin et al. (2011) reported a similarity in gene content but with gene organization that differed from that of Plasmodium spp. (e.g. Feagin et al., 1997) or piroplasms in the genera Babesia or Theileria (see Hikosaka et al. (2010). The mt protein-coding genes of apicomplexan parasites are typically A/T-rich (e.g. ~60-70% A/T). The high A-T bias is indicative of the compositional biases in codon usage. However, this documented bias in codon usage results in few non-synonymous substitutions in the resulting amino acid translation.

2.5.3. The Barcode of Life Initiative

The concept of ‘DNA barcoding’ is a recent molecular innovation. The idea is to design a quick diagnostic genetic method that would function in a way analogous to the 12-digit universal supermarket product code. This molecular method can utilize a standard universal molecular marker present in every species. Hebert et al. (2003a) proposed using a short sequence data from the mitochondrial cytochrome c oxidase I (COI) obtained from different samples as the standard gene for global DNA barcoding for eukaryotes. The premise of this method is based on the concept that, within the COI
gene, genetic variation between species exceeds variation within species (Frezal and Leblois, 2008). Thus, COI or similar suitable DNA sequences from organisms “can be viewed as genetic ‘barcodes’ embedded in every cell” according to Hebert et al. (2004a) that can support identification, delimit species boundaries and catalog taxa, even if they have yet to be described.

DNA barcoding is sequence-based; the four possible bases at each position within a gene provide the possibility of considerable variability that can be exploited. Hebert et al. (2003a) suggested that if the third position nucleotides (i.e. weakly affected by selection) over 15 codons (with 4 possible bases at each of these third positions) were examined; it would generate more than 1 billion unique ‘codes’. This number, the authors argued, will be far greater than what is required to categorize the entire diversity of extant life. Of course, in reality, structural and functional constraints of coding DNA sequences would greatly reduce the number of biologically viable DNA ‘codes’ in 15 consecutive bases in a coding strand of DNA. While a stretch of 45 nucleotides could produce a billion possible codes in principle, Hebert et al. (2003a) argued this concept would be better applied to longer sequences. They calculated that sequences of about 600bp diverging at a rate of 2% per million years (Myr) would have more than 12 diagnostic nucleotide differences between species with a Myr period of reproductive isolation (600×0.02=12). Molecular and archeological records suggest that most species persist for longer than a million years. Hebert et al. (2003b) proposed using a gene with rapidly evolving regions to facilitate species identifications on a scale larger than previously possible using morphological data (Frezal and Leblois, 2008). Additional uses of DNA barcoding methods could be when dealing with new or unknown species (Hebert et al., 2004a; Kumar et al., 2007) or when species–level diagnosis is complex (Wiggins, 1996). This is the case when identifying cryptic species or in cases where it is difficult to associate immature stages to adults stages (Hebert et al., 2004a; Stocke, 2003; Frezal and Leblois, 2008). DNA barcoding could provide data for taxa where morphological visualization is difficult (e.g. midges - Frezal and Leblois, 2008; Ander et al., 2013) or in taxa with pronounced phenotype plasticity (e.g. Alaria species - Lane et al., 2007). Other possibilities may include identifying species from body parts or from a mix of biological specimens (Stocke, 2003; Casiraghi et al., 2010) or when there are
difficulties in matching morphologically divergent male and female members of same species (Hogg et al., 2009).

The single gene, distance measure approach of DNA barcoding for species identification and delimitation has been criticized as inadequate for resolving biological diversity in general and for accurately delimiting the species found within some species complexes in particular (Mayer and Paulay, 2005; Hajibabaei et al., 2006). Other objections in the use of DNA barcoding are the potential for the method to produce misleading results because of the limitation imposed when small taxon samplings or small COI fragment lengths are used for species delimitation (Roe and Sperling, 2007a). Classical identification methods (morphological and molecular) aim to evaluate large numbers of representative samples across wide ecological regions or geographic distance. Use of small numbers of samples or part of a sampled gene may not be enough. Partial COI sequences may prove insufficient for delimiting species in some taxonomic groups. Although Hebert et al. (2003a; 2004b) suggested that examining any stretch of 45 nucleotides could produce up to 1 billion identification labels; DeSalle et al. (2005) noted that such variation is unlikely to produce a viable translated protein in extant species. Therefore, the length of COI sequence required to obtain species-specific sequence (the barcode) reliably is likely to be much longer and required variability may not exist between some closely related taxa. The primary criticism of DNA barcoding by DeSalle et al. (2005) is the almost universal use of computationally simple distance methods (i.e. sequence similarities) and raw distance measures to infer species level groupings. These authors argued that sequences separated by a small genetic distance do not always indicate a close evolutionary relationship because of the lack of a phylogenetic hypothesis in a character-based on distance methods (Koski and Golding, 2001). Character based approaches are more likely to lead to identification of species based on evolutionary descent. Certainly, unlike distance methods, character-based analyses provide the opportunity to test evolutionary hypotheses (DeSalle et al., 2005). Will and Rubinoff (2004) suggested that because DNA barcoding utilizes a single identification method compared to multiple approaches for species delimitation (combinations of character-based schemes, morphological, behavioural and classical taxonomic approaches), the lack of all available data would lead to less logical or robust
conclusions. The DNA barcode method for species identification is therefore less effective than a combined, corroborated, total evidence approach that uses multiple data analyses (Gatesy et al., 2002). DeSalle et al. (2005) suggested a holistic approach that includes morphology, behaviour, biology and life history traits, ecology, and multi-genome genetic data in phylogenetic studies. Cameron et al. (2006) objected to the presumed high costs for DNA barcoding projects compared to currently used methods such as morphological characterization. The various criticisms resulted from the notion that DNA barcoding as proposed by Hebert et al. (2003a) eliminates completely the merit and utility of morphological characters in designating species.

Contrary to this opinion, DNA barcoding is proposed as one application of genetics as an adjunct to the repertoire of classical taxonomic approaches (see Hebert and Gregory, 2005). Classical taxonomic methods could ever only be completed for a tiny percentage of extant species because of the time and costs involved. A better solution may be to acknowledge the limitations of DNA barcoding as outlined by DeSalle et al. (2005) but to utilize the COI sequences provided by DNA barcoding initiatives in character-based analyses. These rapidly growing repositories of accessible sequence data can then act as a starting point for more detailed (if required) species delimitation or evolutionary studies.

2.5.4. GENE OF CHOICE FOR DNA BARCODING

The use of mitochondrial gene sequences to provide data to aid in the identification of species is not entirely new (Williamson et al., 1985; Burgener and Hubner, 1998; Wolf et al., 1999). However, use of a mitochondrial encoded genetic locus as the primary target for species-level analysis of eukaryotes (DNA barcoding) as proposed by Hebert et al., (2003a, 2003b) focused considerable attention on this genome. The selection of the COI locus as a DNA barcode target was simultaneously opportunistic (many COI sequences already existed) and carefully selected (an appropriate rate of genetic divergence for species level analyses with uniparental inheritance). The PCR fragment amplified from the 5´-end of the COI gene is relatively ‘PCR friendly’ with highly conserved regions lying about 600bp apart in many eukaryotes.
A short (~600-800bp) partial COI sequence can still provide the variability needed for distinguishing species (Hebert et al., 2003b; 2004a; 2004b) and differentiating diverse phyla in the animal kingdom (Stoeckle, 2003; Frezal and Leblois, 2008). As a bacterially derived gene, mt COI lacks introns and thus the quality of the sequence is relatively easy to confirm using translation and positional homologies in sequence data are easy to assign. Because it is maternally derived, there is limited recombination, and its haploid mode of inheritance ensures that sequence divergence largely tracks organismal divergence; consequently, COI sequences have considerable utility as a taxonomically informative genetic locus (Hebert et al., 2003a; Roe and Sperling 2007). Highly beneficial in facilitating recovery from suboptimal samples is that the mt genome is abundant in cells; there may be as many as 1000 mitochondria in a eukaryotic cell. Each of those mitochondria may possess 10 mitochondrial genome copies and thus there may be up to 10,000-fold more copies of the mt genome compared to the nuclear genome (Wolf et al., 1999). Consequently, mt COI sequences are comparatively easy to obtain, even from limited genetic samples, and even modest lengths of mt COI sequence frequently contain enough information to discriminate species. This gene has been the most widely used genetic target in DNA barcoding in higher animals; other genetic targets have been found to be more useful for plants (DeSalle et al., 1987; Kjer et al., 2002). The use of mitochondrial SSU rDNA (16S) for taxonomic analyses in protists was complicated by the presence of introns or indels that caused difficulties in sequence alignments (Hebert et al., 2003a). Indels are nucleotide insertions or deletions leading to lengthening or shortening of the sequence. In the protein-coding genes, indels would lead to frame shifts that would make the protein product non-functional and therefore most protein coding genes lack such indels because they are negatively selected strongly.

The bacterially derived mitochondrial cytochrome c oxidase subunit I (COI) gene also benefits from a lack of introns that are found in the intron-exon architecture typical of nuclear genes of eukaryotes (Hogg et al., 2009). Further, the use of COI gene has other important advantages. First, the short segment of the 5’ end of COI gene has conserved regions about 600bp apart making it relatively easy to amplify this target using PCR with conserved primers that can be used with a wide range of taxa (Folmer et
al., 1994; Hebert et al., 2003a; Roe and Sperling, 2007). Secondly, COI sequences possess a nucleotide divergence or evolutionary rate that is compatible with alpha taxonomy (i.e. species identification and differentiation). Hebert et al. (2003a) noted that COI genes like other protein coding genes show high incidence of nucleotide substitutions at the third codon position. The high substitution rate in the third codon position allows for sequence divergence within the COI gene that can be three times greater than mitochondrial 12S or 16S rDNA. Mitochondrial COI still retains comparatively low sequence variation within a single species in most cases (Hogg et al., 2009). Testing the usefulness of DNA sequences derived from the mt COI in seven different animal phyla, Hebert et al. (2003b) obtained COI sequence data that assigned 96% of taxa to their respective phyla when analysed phenetically. Next, they generated COI sequence data for eight orders of insects; the sequences assigned 50 new taxa correctly to their orders. Another set of partial COI sequences for 200 newly analyzed, closely related species (3 allied superfamilies (Geometroidea, Noctuoidea, Sphingoidea) of lepidopterans assigned 150 to species levels. Hebert et al. (2003b) then confirmed the ability of short COI sequences to successfully differentiate taxa previously identified in morphological taxonomic studies. They then used the data as a guide for species diagnoses where there was limited previous taxonomic information. The prospect therefore of using COI barcodes for identification of distantly or closely related species from samples in a variety of conditions (i.e. fresh, stored, whole or in parts) is appealing.

2.5.5. COI Universal Primers

The universal primer set LCO1490 F (5′-GGTCAACAATCATATAAGATATTTG-3′) and HCO2198 R (5′-TAAACTTCAGGGTGACCAAAAAATCA-3′) was demonstrated to amplify a 658 bp fragment of the COI gene in a wide range of invertebrate taxa (Folmer et al., 1994; Hanner, 2005). Consequently, the universal COI primer has proven to be widely applicable to species delimitation and diagnosis for a wide array of animal taxa such as in Table 2.3.
Table 2.3. A selection of published uses of the mitochondrial cytochrome c oxidase subunit I locus for DNA barcoding of metazoan

<table>
<thead>
<tr>
<th>Metazoan Group</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Butterflies</td>
<td>Hebert et al., 2004a</td>
</tr>
<tr>
<td>Birds</td>
<td>Hebert et al., 2004b; Chung et al., 2010</td>
</tr>
<tr>
<td>Spiders</td>
<td>Greenstone et al., 2005</td>
</tr>
<tr>
<td>Mayflies</td>
<td>Ball et al., 2005</td>
</tr>
<tr>
<td>Midge</td>
<td>Carew et al., 2005</td>
</tr>
<tr>
<td>Amphibians</td>
<td>Vences et al., 2005</td>
</tr>
<tr>
<td>Snakes</td>
<td>Pook and McEwing, 2005</td>
</tr>
<tr>
<td>Leeches</td>
<td>Bely and Weisblat, 2006; Oceguera-Figueroa et al., 2010</td>
</tr>
<tr>
<td>Neotropical bats</td>
<td>Clare et al., 2006</td>
</tr>
<tr>
<td>Nymphalids</td>
<td>Hajibabaei et al., 2006</td>
</tr>
<tr>
<td>Free living nematodes &amp; Alaria</td>
<td>Lane et al., 2007</td>
</tr>
<tr>
<td>Ciliates</td>
<td>Chantangsi et al., 2007</td>
</tr>
<tr>
<td>Coral reefs, amphipods &amp; spong</td>
<td>Park et al., 2007</td>
</tr>
<tr>
<td>Fishes, sharks, prepared seafood</td>
<td>Rasmussen and Morrissey, 2008; Bucklin et al., 2011</td>
</tr>
<tr>
<td>Marine metazoa</td>
<td>Steink et al., 2009</td>
</tr>
<tr>
<td>Soil animals</td>
<td>Rougerie et al., 2009</td>
</tr>
<tr>
<td>Equids</td>
<td>Orlando et al., 2009</td>
</tr>
<tr>
<td>Parasitoid wasps</td>
<td>Zaldívar-Riverón et al., 2010</td>
</tr>
<tr>
<td>Opossum</td>
<td>Cervantes et al., 2010</td>
</tr>
<tr>
<td>Tapetinae</td>
<td>Chen et al., 2010</td>
</tr>
<tr>
<td>Ticks</td>
<td>Leo et al., 2010</td>
</tr>
<tr>
<td>Bees</td>
<td>Magnacca et al., 2010</td>
</tr>
<tr>
<td>Parasitoid wasps</td>
<td>Zaldívar-Riverón et al., 2010</td>
</tr>
<tr>
<td>Forensic wildlife cases</td>
<td>Dalton and Kotze, 2011</td>
</tr>
<tr>
<td>Macro-invertebrates</td>
<td>Sweeney et al., 2011</td>
</tr>
<tr>
<td>Primates and other mammals</td>
<td>Luo et al., 2011</td>
</tr>
</tbody>
</table>

The Folmer primers (Folmer et al., 1994) do not work on protists in general and on apicomplexan parasites in particular. As observed in most protistan barcoding projects, amplifying primers designed uniquely for the mt COI gene of the target taxa are required because the universal iBOL (Folmer et al., 1994) primers (Roe and Sperling, 2007a; 2007b) cannot be used. However, the working group of Consortium of the Barcode of Life proposed that the primer design for barcode purposes should flank the position 58–706 region of the mouse mitochondrial COI gene, a position somewhat replicate of the Folmer et al. (1994) region (Hanner, 2005).

2.5.6. COI SEQUENCES IN THE APICOMPLEXA

Use of mt COI sequences for characterization or species delimitation of protists has been limited. Unpublished work (T. Gariepy, S. Shokralla, P.D.N. Hebert and T.R. Gregory, personal communication) documented the use of COI sequence for the identification of protists in a number of groups including ciliates (Barth et al., 2006;
Chantangsi et al., 2007; Struder-Kypke and Lynn, 2010), red algae (Robba et al., 2006; Saunders, 2008), diatoms (Evans et al., 2007) and parasitic coccidia (Ogedengbe et al., 2011). Many researchers attempting to use mt COI sequences for DNA barcoding of protists amplify a portion of the mt COI gene using taxon-specific primers rather than universal primers (Roe and Sperling, 2007b). Ogedengbe et al. (2011) generated DNA barcodes for parasitic coccidia using coccidia-specific primers that amplified ~500 to 800bp of the mt COI gene that overlapped with region of the mt COI amplified by the Folmer universal primers by ~328 bases. Exceptionally, 6 morphospecies of free-living amoebae belonging to the genus *Vannella* were the only protists barcoded successfully using the universal Folmer amplification primers (Nassonova et al., 2010; Hebert et al., 2003a). Mitochondrial COI sequences show strong promise for utility as a diagnostic barcode; such sequences may be combined with sequences from other mitochondrial genes or nuclear genes for a wide variety of protists (http://www.bolnet.ca/rp_protists.php).

Although the COI gene has been used extensively for species identification of a wide variety of organisms, COI sequences have not been used extensively with most apicomplexan parasites due to of a lack of PCR primers that reliably amplify this target. Nonetheless, the COI gene is present in most groups of apicomplexan parasites with the exception of *Cryptosporidium* species and, perhaps, the gregarines. No doubt, mt COI sequences can be used to infer evolutionary developments and relatedness among those members of the Apicomplexa possessing this genetic locus; despite this, DNA barcoding projects of apicomplexan parasites are in their infancy.

Partial COI sequences have been obtained from many haemosporinid species (i.e. *Plasmodium* species - Feagin et al., 1991; Wilson et al., 1991; Feagin, 1992; Preiser et al., 1996), from the piroplasms in the genera *Babesia* and *Theileria* (Kairo et al., 1994; Omori et al., 2007; Lau, 2009; Hikosaka et al., 2011), and, only recently, from a range of coccidia belonging to the families *Eimeriidae* and *Sarcocystidae* (see Ogedengbe et al., 2011). The utility of partial mt COI gene sequences for species delimitation of coccidia (a measure of the likely utility of this locus for reliable species identification) was compared with the utility of the more commonly used nuclear 18S
rDNA sequences for the same purpose. Using over 120 COI sequences from parasites in the genera Caryospora, Eimeria, Cystoisospora, Isospora, Neospora and Toxoplasma, Ogedengbe et al. (2011) were able to show that a 500-800 bp portion of the mt COI gene was more reliable for species delimitation than complete nuclear 18S rDNA sequences. The partial COI sequences had much lower within species sequence diversity compared to the within species variation of complete 18S rDNA sequences. At the same time, the genetic distance between species at the COI locus was much higher than the genetic distance between species using nu 18S rDNA. The more favorable genetic distance ratio (low intraspecific variation and high intraspecific variation) of the COI sequences compared with 18S rDNA makes COI sequences more desirable and reliable genetic markers. Ogedengbe et al. (2011), using partial mt COI sequences from seven Eimeria species infecting chickens and related parasites demonstrated that the shorter COI sequences had much better utility for species delimitation than the longer nu 18S rDNA sequences from the same taxa. Ogedengbe et al. (2011) concluded that comparatively short mt COI sequences from unknown samples were likely to be assigned correctly to COI-characterized species with higher probability than even complete 18S rDNA sequences from the same parasites. For example, the probability of successfully delimiting two closely related Eimeria species, E. tenella and E. necatrix, using COI partial sequences was 97% whereas the probability dropped to only 78% when complete 18S rDNA sequences were used (Ogedengbe et al., 2011). This high probability of correctly assigning unknown samples to COI-characterized species is highly desirable in a genetic locus to be used for molecular identification of species.

2.6. PROBLEMS AND CONCLUDING REMARKS

The evolutionary history of apicomplexan parasites is uncertain and their biodiversity has only started to be explored despite their roles in some of the most serious parasitic diseases of humans and other mammals. As outlined above, morphological and biological characterization of this group has been shown to be inadequate in delimiting species and defining phylogenetic relationships among those species.
Phylogenetic reconstructions of taxa within the phylum Apicomplexa have been based largely on a few, comparatively biologically conserved, members of the phylum (Leander et al., 2003a; Morrison, 2009). As Morrison (2009) noted, taxonomic groups within the Apicomplexa “were originally designed to be utilitarian rather than to reflect evolutionary history”. Consequently, taxa sampling within the Apicomplexa has been limited in a way to a few species within subsets with medical implication to humans rather than strictly for the purpose of phylogenetics. While phylogenetic studies might have been thorough with some taxa for which sufficient taxon sampling was available (notably parasites infecting humans or animals such as the malarial or coccidial parasites), there is a paucity of sampling of other biologically diverse taxa within the phylum. Understandably, the resulting evolutionary hypothesis for members of the phylum is incomplete and probably unreliable because of the lack of intermediate taxa within the phylogeny; the current hypotheses may even be completely incorrect (Morrison, 2009). Limited molecular information has been obtained from parasites such as the so called ‘unimportant’ gregarines or haemogregarines parasites found in hosts examined infrequently and consequently sampled inadequately. Besides, phylogenetic analysis of Apicomplexa has largely been based on apparent morphological selection of ultrastructural synapomorphies of the apical complex, with less reflection on recent molecular data (Siddall, 1995; Morrison, 2009). Phylogenetic reconstruction based on morphology is complicated by the fact that many apicomplexan parasites complete life stages are unknown. Although monophyly of species within the Apicomplexa is supported by synapomorphic ultrastructural characters (e.g. possession of the apical complex), it has not been consistently supported by molecular data. A case in hand is the difficulties of resolving phylogenetic relationships among members of family Sarcocystidae (Toxoplasma, Cystoisospora, Neospora, Hammondia, Sarcocystis, Besnoitia and Frenkelia spp.). Molecular studies and biodiversity studies in last two decades focused on the use of nuclear genetic loci, such as the large and small subunit ribosomal DNA or ribosomal internal transcribed spacer regions (ITS-1 and ITS-2), resulting in many taxonomic reclassifications as more data became available. Phylogenetic relationship of tissue coccidia as deduced based on the nuclear 18rDNA sequences was unresolved for two genera, Sarcocystis and Hammondia species. When
*Frenkelia* species are included in phylogenetic analyses of parasites within the subfamily Sarcocystinae, *Sarcocystis* species do not form a monophyletic clade making the genus *Sarcocystis* paraphyletic. In such analyses, *Sarcocystis* species are found in two large monophyletic clades; the first contains only *Sarcocystis* species and this clade is the sister clade to the second that contains both *Sarcocystis* and *Frenkelia* species. Molecular phylogenetic studies thus far suggest *Frenkelia* should be merged with *Sarcocystis*. Paraphyletic genera are also found in the other subfamily in the Sarcocystidae, the Toxoplasmatinae. *Toxoplasma gondii* was found to form a monophyletic clade with *Hammondia hammondi*; both use felids as definitive hosts. A second monophyletic clade within the Toxoplasmatinae contained *Neospora caninum* with a number of *Hammondia* species that infect canids. This paraphyly of *Hammondia* species demonstrates a conflict between evolutionary relationships inferred from molecular data and taxonomic assignments that were based on morphological and life cycle details (Ellis and Morrison, 1995; Ellis et al., 1998; Mugridge et al., 1999a; 2000).

Adoption of use of mt COI sequences for molecular identification or phylogenetic studies of apicomplexan parasites has been slow. Factors slowing collection of data from this genetic locus for such studies includes a lack of suitable primers, a lack of a characterized barcode region within the gene and, finally, difficulty of obtaining parasite samples free from infected host tissues or cells (Templeton et al., 2010). Progress must be made to identify one or more suitable barcode region(s) and to develop group specific primer sets when a single pair is inapplicable to all taxa within the phylum (T. Gariepy, S. Shokralla, P.D.N. Hebert and T.R. Gregory, personal communication). Accumulation of a large database of apicomplexan mt COI sequences, whether complete or partial, for use as barcode data awaits resolution of these technical issues; even when such hurdles are overcome, some apicomplexan taxa lacking the appropriate mitochondrial genome target (e.g. cryptosporidia and gregarines) require a different genetic locus, perhaps COWP and HSP70 genes.

Complete mitochondrial genome sequences are now available for haemosporids (in *Plasmodium* and related genera), piroplasms (in *Babesia, Theileria* and related genera) and coccidia (*Eimeria* spp.) (see Kairo et al., 1994; Preiser et al., 1996; Feagin et
al., 1991; Hikosaka, 2010). Other apicomplexan parasites such as the adeleid coccidia (e.g. *Hepatozoon* and other haemogregarines), tissue coccidia (e.g. *Sarcocystis*, *Toxoplasma* or *Neospora*) are being targeted for whole mitochondrial genome amplification and sequencing. When generated, this expanded molecular dataset may permit better resolution of the phylogenetic relationships among apicomplexan parasites. Generally, there exist few molecular data derived from adeleid coccidia or gregarines. The gregarines, in particular, are morphologically variable but only a tiny portion of the extant biodiversity of the gregarines has been explored. Interrelationships among described families and genera of gregarines are not known (Leander et al., 2003a). Genome surveys have provided 2657 sequences from *Ascogregarina taiwanensis* (Templeton et al., 2010) and 9301 sequences from *Gregarina niphandroides* available on public databases (e.g. http://www.ncbi.nlm.nih.gov/nuccore, accessed 2015-02-18). No evidence of a mitochondrial genome has been detected in either gregarine genome project. The three typical protein-coding genes found in apicomplexan mt genomes (COI, COIII and CytB) appear to be lacking in gregarines, despite the ultrastructural appearance of what appear to be typical protistan mitochondria in at least some gregarines. Nuclear coded proteins (i.e. heat shock protein 90 (hsp90), cytosolic 70 kDa heat-shock protein (HSP70), dihydrofolate reductase-thymidylate synthase (DHFR-TS), protein kinase (cdc2), beta-tubulin) have, however, been used in phylogenetic analyses of some apicomplexan parasites including gregarines (Zhu et al., 2000; Leander et al., 2003a, Leander and Keeling 2004).

While efforts may be targeted to species of medical and veterinary importance in humans, wildlife, livestock and companion animals (e.g. *Plasmodium*, *Babesia*, *Toxoplasma* and *Eimeria* species) because of their association and impact in disease emergence and zoonoses, attention is required equally for other members of the phylum. This will provide supporting and comparative data that will help resolve phylogenetic issues among pathogenic species. Exploring genome evolution among major groups of apicomplexan parasites (as yet unexplored adeleid coccidia, gregarines and tissue coccidia) might throw light on evolution of the group as a whole and help better define evolutionary relationships within the phylum.
The International Barcode of Life (iBOL) initiative recently set up a new program namely HealthBOL with the mandate to identify and collate about 10,000 COI partial sequences from characterized parasites and vector species through coordinated efforts among public health practitioners, taxonomists and molecular specialists. The specimens and COI sequences generated through this initiative will be linked with a voucher identifier to enhance parasite diagnosis and biomonitoring. The HealthBOL efforts will also aid in identification of parasites where classical methods for identification had failed as well as aid large scale distribution of barcode data (Gariepy et al., unpublished data). The possibility of studying evolutionary history by comparisons of the same gene target across the wide diversity of taxa is a prospect worth investigating. Use of nuclear 18S rDNA sequences to infer relationships among apicomplexan parasites was a good first step but this genetic locus has failed to provide sufficient resolution at the species level to generate a robust phylogenetic reconstruction for the phylum (e.g. Morrison, 2009). Phylogenetic studies on taxa within the phylum Apicomplexa is at a point when both phenotypic (morphological) and genotypic (molecular) information could and should be correctly and gainfully combined (Tautz et al., 2003; Morrison, 2009).

The mt cytochrome c oxidase subunit 1 (mt COI) gene shows some promise as an effective DNA barcode target for those parasites that possess a mitochondrial genome. This thesis is an attempt to confirm the utility of this genetic locus for species delimitation/identification and molecular phylogenetics. In doing so, the research described herein should provide the necessary background information (e.g. group-specific PCR primer sets) and techniques to permit mt COI sequences to be used successfully as a molecular species identification tool (DNA barcode) with most parasites in the phylum Apicomplexa.
3. RESEARCH OBJECTIVES

Research Objective 1:
To determine whether DNA barcodes can be generated from the vast majority of aerobic apicomplexan parasites through representative sampling of a biologically diverse range of apicomplexan parasites.

Research Objective 2:
Devise a robust and simple DNA barcode technique that can be applied successfully to identify coccidia. Such a technique could be applied to long stored (i.e. frozen, air dried or liquid nitrogen stored) fecal /oocyst samples or within preserved or fresh tissues so long as DNA could be obtained from the sample.

Research Objective 3:
Expand the library of COI sequences from various apicomplexan species for iBOL, HealthBOL and GenBank.

Research Objective 4:
To determine if the genetic diversity found within the selected COI region (the DNA barcode region) can be exploited to identify multiple Eimeria species infecting poultry.

Research Objective 5:
Explore the diversity of mt genome structure across the Apicomplexa by generating complete, annotated mt genome sequences from as many biologically diverse apicomplexan parasites from groups that are not already well represented in publically available databases.
4. RESEARCH APPROACHES

Fullfilling the objectives for this study could involve the following:

- Use publically available partial mt COI or complete mt genome sequences to design widely applicable degenerate PCR amplification primers applicable for use with all or most aerobic Apicomplexa;
- Select a suitable DNA barcode region and use PCR and sequencing to generate DNA barcodes from a biologically diverse range of aerobic apicomplexan protozoan parasites of agricultural, companion and wildlife animals. Initially, coccidian parasites of chicken and turkey will be assessed but this will be expanded to explore examples from all major lineages of apicomplexan parasites.
- Generate a robust phylogenetic hypothesis for apicomplexan parasites using partial sequences of the mitochondrial COI gene.
- Use COI sequences generated in this study to assess paraphyly within the family Sarcocystidae and broadly across the Apicomplexa.
- Using *Eimeria* species infecting poultry as a test group, attempt the simultaneous identification of *Eimeria* species within a mixed sample using real-time quantitative PCR with multiple species-specific probes or pyrosequencing of PCR amplification products.
- Amplify mitochondrial whole genomes from eimeriid coccidia, tissue coccidia, gregarines and adeleid coccidia to complement the existing mt genomes available from haemosporinid and piroplasmid parasites. The circular mapping nature of most apicomplexan mt genomes should permit complete mt genome amplification using long-range PCR using ‘outward-facing’ primers designed within partial mt COI sequences or conventional PCR using a small number of overlapping fragments.
- Combine existing and newly generated mt genome sequences to explore the diversity of mt genome organization and content within the phylum Apicomplexa.
5. LIMITATIONS

5.1. CRYPTOSPORIDIUM: AN ALTERNATE BARCODE TARGET REQUIRED

The association of Cryptosporidium spp. and cryptosporidiosis as a complication of human immunodeficiency virus (HIV) infections has generated an increased interest in the parasite. The genus Cryptosporidium was first described by Tyzzer in the early 1900s. Two species were described in mice namely C. muris from the gastric epithelium in 1907 and C. parvum from the intestine in 1912. An additional 24 species have been described subsequently from a wide spectrum of hosts (Ryan et al., 2014). Cryptosporidium hominis (previously termed C. parvum human genotype 1) can be transmitted from animals to humans or from humans to humans (Xiao et al., 2002; 2004), and is water-borne as well (Morgan et al., 1995). Cryptosporidium species are amitochondriate and thus, do not possess the mitochondrial genes nor the mt complexes involved in Kreb’s cycle (Abrahamsen, 2004). Rather, a trans-hydrogenase gene has been identified in this parasite. Alternate DNA barcode targets (see http://www.barcodinglife.org) with sufficient interspecific genetic variable will be required for Cryptosporidium spp. (see Scicluna, 2006). Successful molecular identification and genotyping of Cryptosporidium species have been accomplished using the small subunit rDNA as the universal marker and the Cryptosporidium oocyst wall protein (COWP) for species differentiation. Two other conservative markers, actin and 70 kDa heat shock protein (HSP70) genes have also been used successfully for genotyping of Cryptosporidium species (Šlapeta, 2006). In particular, sequences obtained from the 70kDa heat shock protein were more useful in epidemiological surveys and would likely be useful as a DNA barcode locus for these parasites (Šlapeta, 2006).

5.2. SMALL SAMPLE SIZES

Apicomplexan parasites are unicellular parasites and, although cosmopolitan, they are difficult to isolate (Morrison, 2009). The vast majority of apicomplexan parasites are refractory to axenic growth in vitro; their dependence on the host cells for essential metabolites makes this group of parasites difficult to cultivate in vitro (Doran, 1973) and a challenge with which to work. In sequencing mitochondrial COI sequences to identify
various apicomplexan parasites, sufficient taxon sampling is hampered by the difficulty of isolating or cultivating many of these parasites in vitro.

5.3. TYPE SPECIMENS

According to the International Commission on Zoological Nomenclature (1999), every parasite species should have a designated type specimen that will be an objective standard reference for that parasite species against which all newly described species types will then be compared and named. Most apicomplexan parasites do not have a designated type specimen; preserving coccidial oocysts as permanent type specimens is particularly difficult (Bandoni and Duszynski, 1988). Recently, the ICZN’s approval of phototypes (illustrations or photomicrographs illustrating features of a newly named parasite that can act as the name-bearing type specimen) has addressed this difficulty somewhat. Barcode of Life Initiative is compiling a gallery of micrograph images that will serve as an efficient and effective approach that can help to reduce errors in species identification.
6. SEQUENCING THE COMPLETE MITOCHONDRIAL GENOME OF EIMERIA MITIS STRAIN USDA 50 (APICOMPLEXA: EIMERIIDAE) SUGGESTS CONSERVED START POSITIONS FOR MTCOI-AND MTCOIII-CODING REGIONS

(The contents of this chapter have been published as follows: M.E. Ogedengbe, M. A. Hafeez, J. R. Barta, 2013. Sequencing the complete mitochondrial genome of Eimeria mitis strain USDA 50 (Apicomplexa: Eimeriidae) suggests conserved start positions for mtCOI-and mtCOIII-coding regions Parasitology Research 112: 4129-4136.

6.1. ABSTRACT

Four complete mitochondrial (mt) sequences from a single-oocyst-derived line of Eimeria mitis USDA 50 were obtained (three from cloned whole-genome PCR products, one from directly sequenced whole-genome PCR product). The mt genome is 6,408 bp long with three genes (CytB, cytochrome c oxidase subunit I (COI) and cytochrome c oxidase subunit III (COIII)) and many rDNA fragments (large subunit rDNA 13, small subunit rDNA 10); organisation was identical to other Eimeria sp. mt genomes. Conserved start codon positions for both COI and COIII are suggested for all Eimeria mt genomes; these start codon positions exist and may also be conserved, in related apicomplexan parasites. Within the three separate cloned PCR products of near-complete mt genomes, there were 26 nucleotide differences (collectively) compared to the directly sequenced mt genome. These changes appear to be base misincorporations during PCR. Direct sequencing of long PCR amplification products may be more likely to generate accurate mt genomic sequences than cloning and subsequent sequencing.

6.2. INTRODUCTION

Coccidiosis occurs wherever chickens are reared (Dalloul and Lillehoj 2006). The disease frequently presents as simultaneous infections with multiple Eimeria species (Apicomplexa: Eimeriidae). The biology and relationships have been described for the Eimeria species involved in the disease: Eimeria tenella, Eimeria necatrix, Eimeria maxima, Eimeria praecox, Eimeria mitis, Eimeria brunetti and Eimeria acervulina (Shirley et al. 1983; Reid and Long 1979; Costa et al. 2001; Barta et al 1997). Global economic impact of coccidiosis in chickens is estimated at US$2.4–3 billion per annum (McDougald and Reid 1997; Fernandez et al. 2003b; Bera et al 2010).
Like many members of the phylum Apicomplexa, *Eimeria* species possess a nuclear genome and two extra nuclear genomes: ~35-kilobase (kb) apicoplast genome (non-photosynthetic plastid) and ~6-kb mitochondrial (mt) genome (Wilson and Williamson 1997; Feagin 1994, 2000). Although mt functionality is widely conserved, gene content, size, arrangement and gene expression vary dramatically among eukaryotes (Gray et al. 1999, 2001). Mitochondrial genome size ranges from more than 360 kb in the flowering plant Arabidopsis to only ~6 kb in *Eimeria* and related apicomplexan parasites (Gray et al. 1999). The structure and gene arrangement of apicomplexan mt genomes is highly variable. Apicomplexan mt genomes have been recorded as linear concatemers of a basic mt genomic unit as in *Plasmodium* (see Preiser et al. 1996) and *Eimeria* species (Hikosaka et al. 2011), or as linear genomes possessing terminal inverted repeats reminiscent of telomeric structures (Kairo et al. 1994; Wilson and Williamson 1997). In *Eimeria* species and other apicomplexan parasites, mt genomes possess coding regions for only three proteins: cytochrome c oxidase subunit I (COI), cytochrome c oxidase subunit III (COIII) and cytochrome b (CytB). In addition to the three retained protein-coding sequences, numerous fragments of small subunit (SSU) and large subunit (LSU) rRNA genes exist but apicomplexan mt genomes lack intact 5S rRNA and tRNA (Gray et al. 2004; Hikosaka et al. 2011; Lin et al. 2011).

Complete mt genome sequences have been published for the seven commonly recognized *Eimeria* species that infect chickens (Hikosaka et al. 2011; Lin et al. 2011; Liu et al. 2012). In the present study, we report on the mt genome sequence from a single-oocyst-derived clonal line of *E. mitis* USDA 50. We generated one complete mitochondrial genome sequence based on PCR amplification followed by direct sequencing, as well as three additional complete mitochondrial genome sequences based on clones of PCR product amplified from the same template. Divergence was observed among all sequenced clones of the *E. mitis* mt genome, but the observed polymorphisms were not reflected in sequence obtained from directly sequenced PCR product from the same template. These variations were likely base misincorporation artefacts that arose during PCR amplification, suggesting that sequences generated from cloned mt genomes should be used with caution. Finally, examination of the available open reading frames
(ORF) for three protein-coding mt genes suggested that some revisions to the current gene annotations of *Eimeria* sp. mt genomes are warranted.

6.3. MATERIALS AND METHODS

6.3.1. PARASITES AND DNA EXTRACTION

The USDA 50 strain of *E. mitis* Tyzzer 1929 was propagated in coccidia-free chickens in the Central Animal Facility's Isolation Unit (University of Guelph, ON, Canada). Chickens were provided with feed and water ad libitum; all experimental manipulations were reviewed and approved by the University of Guelph's Animal Care Committee and complied with the Canadian Council on Animal Care's ‘Guide to the Care and Use of Experimental Animals, vol. 1, 2nd edition’.

A single-oocyst-derived line of *E. mitis* USDA 50 was obtained essentially as described by Remmler and McGregor (1964) with the modification that agar plugs with a single oocyst were delivered individually to the crop of birds within gelatin capsules. Oocysts resulting from single-oocyst isolation were propagated as described by Ogedengbe et al. (2011). Faecal material was homogenized in saturated NaCl (aqueous) and then passed once through a course strainer to remove large debris. The oocysts were partially isolated from faecal debris by flotation using saturated NaCl solution (aqueous) and centrifugation at 2,000×g for 10 min according to the method of Ryley et al. (1976). The floating oocysts in saturated NaCl were decanted from the top of the centrifuge containers and diluted at least 10× in distilled water prior to a second centrifugation at 2,000×g for 10 min to pellet the oocysts. The supernatant was decanted; the partially purified oocysts were suspended in 2.5 % potassium dichromate (w/v, aqueous) and then sporulated on a rotary shaker at 26°C for at least 72 h. Fully sporulated oocysts were stored in 2.5 % potassium dichromate solution at 4 °C. Oocysts were further separated from faecal debris by repeating the flotation method (Ryley et al. 1996). Finally, the partially purified sporulated oocysts were collected by centrifugation and then surface-treated using household bleach (4.25 % sodium hypochlorite w/v, aqueous) for 10 min on ice to degrade exogenous agents or DNA prior to DNA isolation from the parasites within the oocysts.
Extraction of DNA was accomplished from the surface-treated (bleached) sporulated oocysts using DNAzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions with the following modification: Once oocysts were suspended in DNAzol reagent, 0.8 g of 0.5 mm-diameter glass beads (Ferro microbeads; Cataphote Division, Jackson, MS, USA) was added to the oocyst/DNAzol mixture. These parasites and glass beads in DNAzol were then vortexed for 30–60 s to aid in oocyst disruption and release of parasite DNA from the oocysts. After vortexing and incubation, the DNAzol solution containing disrupted parasites and extracted DNA was transferred to a new container and the rest of the DNA isolation was performed according to the manufacturer's instructions. The quantity and purity of DNA was assessed spectrophotometrically using a NanoDrop 2000 instrument (NanoDrop Products, Wilmington, DE, USA).

6.3.2. PCR Amplification

A long PCR reaction was performed using the QIAGEN LongRange PCR Kit (QIAGEN, Valencia, CA, USA) according to the manufacturer's instructions. *E. mitis* DNA (100 ng) was used as the template in a PCR reaction containing 1× PCR buffer, 4 mM MgCl2, 2.5 units Platinum Taq (Invitrogen) and 10 pM of the ‘*E. mitis*-WG-MT-Inv F’ primer (5′-ACTGGATCACCATTAAAGGCAGCA-3′) and the ‘*E. mitis*-WG-MT-Inv R’ primer (5′-TCTGGTTCTTTTG GACACCCAG-3′). These primers were designed to anneal 17 bp apart near the middle of the published partial sequence for mitochondrial cytochrome *c* oxidase subunit I gene of *E. mitis* USDA 50 (bases 326–303 and 344–364, respectively, in GenBank sequence HM771681).

The PCR reaction was run using an MJ Mini thermal cycler (Bio-Rad, CA, USA) with the following cycling conditions: initial denaturing at 93 °C for 3 min; 35 cycles of 93 °C for 15 s, 50 °C for 30 s, 68 °C for 6 min; and a final extension cycle of 68 °C for 10 min. A portion of the PCR reaction was electrophoresed using a 0.8 % agarose gel at 50 V using crystal violet as the in-gel DNA stain (Yang et al. 2001; Turgut-Balik et al. 2005) and compared to the DNA size ladder (Lambda DNA/HindIII Marker, Thermo Fisher Scientific, Pittsburgh, PA, USA) to identify the band of interest. DNA bands were
excised from the gel and then purified using a QIAGEN gel extraction and purification kit.

6.3.4. CLONING

Gel-purified PCR products were cloned using a TOPO XL Cloning kit with One Shot TOP10 Chemically Competent E. coli kit (Invitrogen) according to the manufacturer's instructions. Two methods were used to assess clone sizes for selection of positive clones (Campbell and Choy 2001). Selected positive clones were propagated in Luria broth (LB) at 37 °C overnight and purified with PureLink™ Quick Plasmid Miniprep Kit (Invitrogen).

6.3.5. SEQUENCING

Sequencing was performed using a primer walking strategy in both directions until each near-complete mitochondrial genome length clone or PCR product was sequenced completely in both directions. Sequencing was conducted by the Molecular Biology Unit of the Laboratory Services Division, University of Guelph (Guelph, ON, Canada) using an ABI PRISM 7000 Sequence Detection System (Applied Biosystems Inc., Foster City, CA, USA). For comparison, three different clones were selected randomly from among 35 clones that were confirmed to have the correct insert length of ~6 kb. Sequencing primers and their locations in the mitochondrial genome of E. mitis are presented in Table 6.1.

6.3.6. SEQUENCE DATA ASSEMBLY

Sequencing chromatograms were trimmed and assembled using the de novo sequence assembler within the bioinformatics software package Geneious version 6.1 (and later versions, available from http://www.geneious.com/) to provide the primary sequence of the near-complete mt genome. Multiple sequences from PCR products targeting the COI gene of the E. mitis USDA 50 mitochondrion in the region that overlapped the inverse PCR primer-binding sites were obtained; all were identical in sequence through both primer-binding sites and the intervening 17 bp of the COI-coding sequence (CDS) (data not shown). For this reason, the 62 bp of this region was used to connect the inverse PCR products and complete these mitochondrial genome sequences for E. mitis USDA 50.
### Table 6.1. Sequencing primers and their locations in the mitochondrial genome of *Eimeria mitis* USDA 50 Clone 24 (KC409031)

<table>
<thead>
<tr>
<th>Sequencing Primer</th>
<th>Genome Location</th>
<th>Gene or Region</th>
<th>Primer Sequence (5´-3´)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Em-WG-MT_344R</td>
<td>344 → 323</td>
<td>CytB Gene</td>
<td>GTAGGAATCCTRAATTCCCAACC</td>
</tr>
<tr>
<td>Em-WG-MT_710F</td>
<td>710 → 729</td>
<td></td>
<td>CCATTAGGTACAGAAAACTGC</td>
</tr>
<tr>
<td>Em-WG-MT_1308R</td>
<td>1,308 → 1,286</td>
<td>COI Gene</td>
<td>GCAAACCATAYATAGTAAATACC</td>
</tr>
<tr>
<td>COI-400F</td>
<td>1,610 → 1,629</td>
<td></td>
<td>GGDCAGTTTGGTCGACGAC</td>
</tr>
<tr>
<td>Em-WG-MT InvF</td>
<td>1,935 → 1,912</td>
<td>COI Gene</td>
<td>ACTGGATCACCATAAAGGCAGCA</td>
</tr>
<tr>
<td>COI-1202R</td>
<td>2,412 → 2,391</td>
<td></td>
<td>CCAAKRAYHGCACCAAGAGATA</td>
</tr>
<tr>
<td>Em-WG-MT_4506F</td>
<td>4,501 → 4,522</td>
<td>COIII Gene</td>
<td>GCTATTAGTACTGAGTTACTAC</td>
</tr>
<tr>
<td>Em-WG-MT_4993R</td>
<td>4,984 → 4,964</td>
<td></td>
<td>GAGGAATACCATAAATAGTAG</td>
</tr>
<tr>
<td>Em-WG-MT_2768F</td>
<td>2,767 → 2,786</td>
<td></td>
<td>GGTCAGCCGGGATAACAGG</td>
</tr>
<tr>
<td>Em-WG-MT_3658F</td>
<td>3,652 → 3,671</td>
<td></td>
<td>CTGCCGAGAAGGAGGATG</td>
</tr>
<tr>
<td>Em-WG-MT_4072R</td>
<td>4,067 → 4,048</td>
<td>Fragmented rDNA</td>
<td>GGTTGGTTCATCTCGACTC</td>
</tr>
<tr>
<td>Em-WG-MT_5416F</td>
<td>5,406 → 5,427</td>
<td>or unassigned</td>
<td>GGTCGAGATAAGCGATCTCATG</td>
</tr>
<tr>
<td>Em-WG-MT_5811R</td>
<td>5,885 → 5,905</td>
<td></td>
<td>GTCGTTACCATACATCGAG</td>
</tr>
<tr>
<td>Em-WG-MT_6219F</td>
<td>6,207 → 6,226</td>
<td></td>
<td>GCATCCATCTACGTCG</td>
</tr>
</tbody>
</table>

Open reading frames were determined using Geneious 6.1 using translation table 4 (i.e. mold/protozoan mitochondrial codon translations); probable CDSs were identified and translated into amino acids using the same translation table, and the resulting hypothetical polypeptide sequence was searched against the non-redundant sequence database using blastp to identify the proteins produced by each CDS.

Putative rDNA fragments were mapped through comparison with mt genomes of *E. tenella* (AB564272, annotated by Hikosaka et al. 2011) and *Plasmodium falciparum* (M76611, functionally annotated, see Feagin et al. 2012).

All available mt genome sequences from *Eimeria* spp. (*E. acervulina* [HQ702479], *E. brunetti* [HQ702480], *E. maxima* [HQ702481], *E. mitis* [JN864949], *E. necatrix* [HQ702482], *E. praecox* [HQ702483], *E. tenella* [AB564272; HQ702484]) were aligned using the multiple alignment algorithm in Geneious 6.1. The newly obtained PCR-based mt genome sequence and three cloned mt genome sequences obtained were linearized at the same position so that all genome sequences had the small subunit rDNA ‘A’ fragment (ssuA) starting at position 1 to permit alignment with related sequences obtained from GenBank. Start and stop codon positions for the identified CDS were compared with the other aligned *Eimeria* sp. mt genome sequences.
6.4. RESULTS AND DISCUSSION

The complete mitochondrial genome of the North American strain *E. mitis* (USDA 50) based on direct PCR product sequencing had a unit length of 6,408 bp (accession number KF501573). Mitochondrial genomes for three clones of *E. mitis* had unit lengths of 6,407 bp (clone 18—KC409029) or 6,408 bp (clone 20—KC409030, clone 24—KC409031). A high A/T base composition bias was noted (67.40 %). Thymine (36.4 %) was the most common nucleotide, followed by adenine 30.9 %, while cytosine and guanine bases were 16.3 and 16.4 %, respectively.

**Table 6.2.** Pairwise sequence identity (upper half) and number of nucleotide differences (lower half) among four complete mitochondrial genome sequences of *Eimeria mitis*.

<table>
<thead>
<tr>
<th></th>
<th><em>Eimeria mitis</em> USDA 50 PCR-based, KF501573</th>
<th><em>Eimeria mitis</em> USDA 50 Clone 18, KC409029</th>
<th><em>Eimeria mitis</em> USDA 50 Clone 20, KC409030</th>
<th><em>Eimeria mitis</em> USDA 50 Clone 24, KC409031</th>
<th><em>Eimeria mitis</em> China Strain JN864949</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Eimeria mitis</em> USDA 50 PCR-based, KF501573</td>
<td>99.9%</td>
<td>99.9%</td>
<td>99.9%</td>
<td>99.5%</td>
<td></td>
</tr>
<tr>
<td><em>Eimeria mitis</em> USDA 50 Clone 18, KC409029</td>
<td>7</td>
<td>99.8%</td>
<td>99.8%</td>
<td>99.4%</td>
<td></td>
</tr>
<tr>
<td><em>Eimeria mitis</em> USDA 50 Clone 20, KC409030</td>
<td>7</td>
<td>14</td>
<td>99.8%</td>
<td>99.4%</td>
<td></td>
</tr>
<tr>
<td><em>Eimeria mitis</em> USDA 50 Clone 24, KC409031</td>
<td>8</td>
<td>15</td>
<td>15</td>
<td>99.4%</td>
<td></td>
</tr>
<tr>
<td><em>Eimeria mitis</em> China Strain JN864949</td>
<td>31</td>
<td>38</td>
<td>38</td>
<td>39</td>
<td></td>
</tr>
</tbody>
</table>

Genome organisation consisted of three protein-coding genes (COI, COIII and CytB) and numerous regions encoding rDNA fragments including 13 LSU fragments and 10 SSU fragments (see Figure 6.1). Ribosomal fragments LSU/15(RNA6), SSU/9(RNA5), SSU/1 (RNA14) and RNA7 were annotated in the present *E. mitis* mt genomic sequences and, although not annotated in GenBank, are present in all other *Eimeria* species for which mitochondrial genome sequences are available (data not shown).
6.4.2 START CODON DETERMINATIONS FOR COI, COIII AND CYTB.

For all four mt genome sequences of *E. mitis* USDA 50 obtained in the present study, the start codon for the CytB CDS was an ATG codon located at position 128 of each sequence. This ATG start codon and a poly-T-rich ‘GGTTATGTTTA’ motif just upstream of the start codon were conserved in all *E. mitis* mt genome sequences as well as in all other *Eimeria* species available (data not shown). A single stop codon, TAA, was used to terminate the CDS for CytB, COI and COIII.

The selection of a start codon for the COI-and COIII-coding regions was less obvious when alignments in the area of the start codons were compared among all available mt genome sequences from seven *Eimeria* species. For the COI CDS, various start codon locations have been suggested for the eight previously published *Eimeria* sp. mt genome sequences but the start codon location has been left unresolved for four *Eimeria* species (Figure 6.2, alignment A). The *E. mitis* USDA 50 strain mt genome has a single nucleotide polymorphism (SNP) that eliminates the start codon ATT previous annotated by Liu et al. (2012) for the *E. mitis* China strain (JN864949); the next start codon available in the five *E. mitis* sequences is a TTA start codon located 21 nucleotides (nts) (seven amino acids (aa)) downstream. Although found in all *E. mitis* sequences, this TTA start codon is not found in all of the other avian *Eimeria* spp. Just six nts further downstream is a conserved location for a start codon (with a consensus...
sequence of ATD [i.e. ATG, ATT or ATA]) found in all 12 *Eimeria* sp. mt genome sequences available for study. We therefore conclude that the most likely start codon for all *Eimeria* spp. is this positionally conserved ATD start codon as illustrated in Figure 6.2 (alignment B).

For COIII, the situation regarding start codon position was similar (Figure 6.3, alignment A). An indel in the *E. mitis* USDA 50 sequence eliminated the previously annotated start codon for *E. mitis* China strain (JN864949). In addition, all previously annotated start codons for COIII were located within the highly conserved LSU/1 (LSUA) rDNA fragment and coded for translation products that varied in length (259 or 260 aa). Selection of a start codon position that is conserved among all 12 *Eimeria* sp. mt genome sequences that does not overlap the LSU/1 (LSUA) region is possible by utilizing the conserved TTA start codon located 16–20 nts downstream of this LSU/1 (LSUA) region; use of this conserved start codon produces a COIII product of 251 aa (Figure 6.3, alignment B). Finally, the variability in the length of the sequences between the end of the LSU/1 (LSUA) region and the proposed conserved start codon position suggests that this is not part of the COIII CDS. For both COI and COIII, there is a poly-A-and poly-T-rich region just upstream of the newly proposed start codon locations that may act as a signal motif.

### 6.4.3. Comparison of Cloned versus Directly Sequenced PCR Products.

Collectively, there were 21 SNPs observed between the cloned and sequenced PCR products versus the directly sequenced PCR product; no SNPs were shared between clones at any variable positions. Transitions (19 changes) greatly outnumbered transversions (two changes, one each of T→G and T→A) among the nt substitutions found in the cloned sequences. Type 1 transitions (T→C [eight positions]/A→G [six positions]) were more common than type 2 transitions (C→T [three positions]/G→A [two positions]); the type 1 to type 2 misincorporation ratio was ~3:1. Positional bias of nt substitutions (i.e. preference for any of the first, second or third codon positions) was not observed.
Pairwise alignment of *E. mitis* strain USDA 50 mt genome sequence and that of the *E. mitis* Chinese strain (JN864949) demonstrated 26 nucleotide differences. CytB had five SNPs, of which four were non-synonymous. COI had four SNPs, of which three were non-synonymous, and only a single SNP, a synonymous change, was observed in COIII. Additional SNPs were observed within ribosomal (six within various LSU rDNA fragments) and 10 SNPs (including five apparent indels) were found at unassigned, possibly intergenic, regions.

6.5. DISCUSSION

The complete mitochondrial genome of *E. mitis* USDA 50 was comparable in length (6408 bp) to the mitochondrial genome sequence from a Chinese isolate of *E. mitis* (JN864949) reported by Liu et al. (2012). The newly reported mitochondrial genome sequence was identical at 6,378 sites with a mean pairwise identity of 99.5 % to the previously reported mitochondrial genome sequence for *E. mitis* (Table 6.2, Figure 6.1).

Gene number, annotation and organisation for the sequenced *E. mitis* USDA 50 mt genome in this study were consistent with all *Eimeria* spp. for which mitochondrial genome sequences are available (Hikosaka et al. 2011; Lin et al. 2011; Liu et al. 2012) but varied from those of other apicomplexan parasites (Feagin 2000; Gray et al. 2004; Hikosaka et al. 2010).

The coding region and start codon (ATG) for the CytB gene were annotated similarly in the previously published *E. mitis* genome and seven available mitochondrial genomes from other *Eimeria* spp. (Hikosaka et al. 2010; Lin et al. 2011; Liu et al. 2012). In contrast, our annotation of the location of the start codon for the COIII-coding region in the PCR-based mt genome differed from annotations for other *Eimeria* spp., including *E. mitis* as reported by Liu et al. (2012), because of an indel event (see Figure 6.3) just downstream of previously annotated start of the COIII CDS.

The sequencing chromatograms were examined carefully in this region and the insertion of adenosine in the PCR-based mt genome sequence and all three cloned mt genomes for *E. mitis* USDA 50 was consistent; this insert does not appear to be a sequencing artefact.
In all previous annotations of mitochondrial genomes from *Eimeria* spp. (Hikosaka et al. 2010; Lin et al. 2011; Liu et al. 2012), ATT was the presumed start codon (located eight nts within the LSUA rDNA fragment) for the COIII CDS resulting in translation products 260 to 261 aa in length. This start codon would not be functional in the newly obtained mt genome sequence, and the first available start codon (again ATT) located at nt 4294 would produce a COIII product that was 258 aa long. However, this potential start codon does not exist in any of the other mt genome sequences from other *Eimeria* spp. There is a highly conserved TTA start codon located at position 4310 that is found in all 12 available *Eimeria* sp. mitochondrial genomes. Using this start codon would result in COIII products from all genomes that were exactly 252 aa in length, avoiding the sequence variability otherwise found at the amino-terminal end of the resultant COIII translation products (see Figure 6.3). Interestingly, a start codon (frequently TTA) is located at the same site in the mitochondrial genomes of *P. falciparum* and related apicomplexan parasites, suggesting that perhaps this is the start codon for COIII translation products for all apicomplexan mitochondrial genomes. Functional annotation of mt gene products would be useful for confirming this proposed start codon location.

A similar situation exists in annotating the initiation site for the COI-coding region. If the first potential start codon following or even including part of the stop codon for the CytB gene is used, COI translation products ranging from 480 to 492 aa in length would be produced from the 12 *Eimeria* sp. mt genomes available. However, as the case with COIII, there was a potential start codon (ATG) located at position 1244 of the PCR-based genome sequence (KF501573); putative start codons (ATG or ATA or ATT) are found at the same location in all mitochondrial genomes available for study from *Eimeria* spp., and initiating translation at this codon results in COI products of exactly 480 aa in all cases (see Figure 6.2). In all *Eimeria* spp. this putative translation start codon is located 27 nts upstream from nucleotides encoding the highly con-served ‘Asn–His–Lys’ motif associated with the start of the heme–copper oxidase subunit I core region of COI. This same ‘Asn–His–Lys’ motif is found in the COI translation product from malarial parasites (i.e. 30 nts downstream of a potential start codon at nt 2047 [TTA] in *P. falciparum* [M76611]) as well as in the dinoflagellates (i.e. 30 nts
**Figure 6.2.** Possible COIII start codon positions for 11 *Eimeria* sp. mitochondrial sequences. Alignment A indicates the start codons (where identified) as annotated in GenBank resulting in translation products of 260 to 261aa in length. Alignment B provides an alternative, but conserved, start codon (TTA) identified as present in all species that would produce COIII products that were 258aa long in all cases.

Downstream of a start codon at nt 25 [TTA] in *Alexandrium catenella* [AB374235]). In all *Eimeria* spp., *P. falciparum* and *A. catenella*, there is an A/T-rich region just upstream from the putative start codon for COI that may act as a signal motif (sensu Feagin et al. 1992—signal motif copy 1 at position nts 2036–2050 of M76611). Interestingly, a putative start codon is found just downstream of the 5′ end of the mRNA for COI from *Plasmodium yoelii* (see Suplick et al. 1990) following a short ‘ATTTTTTGTTT’ motif that was found at the 5′ end of both the COI and CytB mRNAs from this parasite. In *P. yoelii*, the start codon for CytB is found 13 nts from the 5′ end of the mRNA following the ‘ATTTTTTGTTT’ motif (Suplick et al. 1990); a putative
start codon for the COI mRNA of the same parasite is found 11 nts from the start of the COI mRNA ‘ATTTTTTGTTT’ motif (the last base of the motif is included in the putative TTA start codon for COI). We suggest that this positionally conserved start codon upstream of codons encoding the highly conserved ‘Asn–His–Lys’ motif may represent the actual start of the coding region for COI in the Apicomplexa (Figure 6.2) regardless of the transcription initiation site.

**Alignment A**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Alignment B**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 6.3.** Possible COI start codon positions for 11 Eimeria sp. mitochondrial sequences. Alignment A indicates the start codons (none identified in 4 of the genome sequences) as annotated in GenBank resulting in translation products of 480 to 492aa in length. Alignment B illustrates that putative start codon (ATG or ATA or ATT) is found at the same location in all Eimeria spp. mitochondrial genomes available for study; initiating translation at this codon results in COI products of exactly 480 aa.
Both *E. mitis* mt genome sequences (Liu et al. 2012 and present study) were generated by PCR amplification and direct sequencing of the resulting PCR product. Initially, during the present study, cloning of near-complete mt genome PCR products was pursued for the purpose of obtaining multiple independent mt genome sequences in an attempt to enhance sequence accuracy. However, there were 21 SNPs noted among the first three clones that were sequenced completely and none of these occurred at the same location on the separate clones. Analysis of the bias in nucleotide substitutions (highly biased transition-to-transversion ratio [19:2] and a high type 1-to-type 2 transition ratio [~3:1]) suggests that these SNPs were simply the result of nucleotide misincorporation during PCR amplification (cf. Olivieri et al. 2010). This observation suggests that direct sequencing is a better approach for obtaining sequence from long PCR products such as the near-complete mt genomes examined in the present study.

Variation can still be expected to occur within a single clonal parasite line (i.e. mitochondrial genome copy-to-copy variation within a single parasite) and possibly among different parasite samples of the same species. *P. falciparum* has had the most numerous complete mt genomes sequenced from any single apicomplexan species. An assessment of over 100 complete *P. falciparum* mt genome sequences yielded 20 positions of variability scattered within the protein-coding genes (five silent and three non-silent changes in COIII, four silent and two non-silent changes in COI and six silent changes only in CytB). This variability was comparable to the 10 SNPs observed between the CDS regions of the two *E. mitis* mt genome sequences now available; this level of variability may reflect strain-level variation in this species. Such strain-level variation was not evident between two isolates of *E. tenella* that had 100 % mt genome sequence identity (AB564272 by Hikosaka et al. 2010, HQ702484 by Lin et al. 2011).

For all *Eimeria* species, the fragmented rDNA regions were more highly conserved than the protein-coding regions due to functional constraints in the former. Twenty-three rDNA fragments were mapped (LSU rDNA—14 fragments, SSU rDNA—8 fragments, unassigned rDNA—1 fragment) as short sequences of varying lengths (16—188 nts); the majority (14) were located between the COI and COIII genes. In addition to previously annotated rDNA fragments (Hikosaka et al. 2011; Liu et al. 2012; Lin et
al. 2011), we were able to map four additional rDNA fragments corresponding to LSU/15(RNA6), SSU/9(RNA5), SSU/1 (RNA14) and RNA7 of *P. falciparum* (M76611) following the naming convention of Feagin et al. (2012). Additional highly conserved regions of the mitochondrial genome of the 12 *Eimeria* spp. remain without annotation and may represent as yet unidentified rDNA fragments. Occurrence of fragmented and incomplete rRNA genes is typical for apicomplexan parasites; fragmented but likely functional rRNA genes have been reported in all apicomplexan mitochondrial genomes studied thus far (Feagin 2000; Feagin et al. 2012; Hikosaka et al. 2010, 2011; Lin et al. 2011).
7. COMPLETE MITOCHONDRIAL GENOME SEQUENCES FROM FIVE EIMERIA SPECIES (APICOMPLEX; COCCIDIA; EIMERIIAE) INFECTION DOMESTIC TURKEYS

(The contents of this chapter have been published as follows: Mosun E. Ogedengbe, Shiem El-Sherry, Julia Whale and John R. Barta, 2014. Complete mitochondrial genome sequences from five Eimeria species (Apicomplexa; Coccidia; Eimeriidae) infecting domestic turkeys. Parasites & Vectors 7:335, http://www.parasitesandvectors.com/content/7/1/335)

7.1. ABSTRACT

Clinical and subclinical coccidiosis is cosmopolitan and inflicts significant losses to the poultry industry globally. Seven named Eimeria species are responsible for coccidiosis in turkeys: *Eimeria dispersa*; *Eimeria meleagrimitis*; *Eimeria gallopavonis*; *Eimeria meleagridis*; *Eimeria adenoeides*; *Eimeria innocua*; and, *Eimeria subrotunda*. Although attempts have been made to characterize these parasites molecularly at the nuclear 18S rDNA and ITS loci, the maternally-derived and mitotically replicating mitochondrial genome may be more suited for species level molecular work; however, only limited sequence data are available for Eimeria spp. infecting turkeys. The purpose of this study was to sequence and annotate the complete mitochondrial genomes from 5 Eimeria species that commonly infect the domestic turkey (*Meleagris gallopavo*).

7.2. INTRODUCTION

As many as seven Eimeria species, *Eimeria dispersa*, *Eimeria meleagrimitis*, *Eimeria gallopavonis*, *Eimeria meleagridis*, *Eimeria adenoeides*, *Eimeria innocua* and *Eimeria subrotunda*, can cause coccidiosis in the turkey, *Meleagris gallopavo* (McDougald, 2003). Coccidiosis is widespread and pathogenic with considerable economic losses to the poultry industry (Chapman, 2008; Dezfolian et al., 2010). These parasites possess morphotypes of oocysts with overlapping biological features that make identification, characterization and diagnosis challenging (Long et al, 1977; Chapman, 2008). Delimiting individual species using morphological features, even when supplemented by 18S rDNA or internal transcribed spacer (ITS) sequence data, has been reported to be less than ideal for coccidia, especially for closely related parasites (Morrison et al., 2004; Rampin et al., 2006; Cook et al., 2010; Poplstein and Vrba, 2011; Ogedengbe et al., 2011; El-Sherry et al., 2013). Sequences from the mitochondrial cytochrome c oxidase subunit I gene (mtCOI) have been shown to be reliable for
delimiting closely related species (Ogedengbe et al., 2011) and the mtCOI locus appears to lack paralog issues associated with rDNA of these parasites (El-Sherry et al., 2013).

A single, complete mitochondrial (mt) genome copy for parasites within the Apicomplexa is about 6KB long (Feagin, 1994; 2000). Genome organisation varies considerably among eukaryotes in general and also within the Apicomplexa (Gray et al., 1999; 2001). Among apicomplexan parasites, genome structures that have been reported include linear concatemers (Preiser et al., 1996; He et al., 2014) linear genomes with terminal inverted telomeric repeats (Hikosaka et al., 2010; Wilson et al., 1997) and circular genomes (Feagin, 2000; Feagin et al., 2012) Regardless of overall genome structure, all apicomplexan mt genomes examined to date possess three genes encoding cytochrome c oxidase subunit I (COI), cytochrome c oxidase subunit III (COIII) and cytochrome b (CytB), as well as numerous fragments of discontinuous and scrambled small subunits (SSU) and large subunit (LSU) rDNA. The specific LSU and SSU rDNA fragments found in the mt genome of apicomplexan parasites differ among distantly related parasites. Unlike many eukaryotic mt genomes, apicomplexan mt genomes do not encode 5S rRNA or tRNAs (Gray et al., 2004; Omori et al., 2007; Hikosaka et al., 2011; Lin et al., 2011; Ogedengbe et al., 2013).

In the present study we report six new PCR generated, complete mt genome sequences from single oocyst derived lines of five Eimeria species infecting turkeys: Eimeria dispersa Briston strain, Eimeria meleagrimitis USMN08-01 strain, Eimeria gallopavonis Weybridge strain, Eimeria gallopavonis USKS06-01 strain, Eimeria meleagridis USAR97 01 strain and Eimeria adenoeides Guelph strain.

7.3. METHODS

7.3.1. PARASITES

Six single oocyst-derived lines of five Eimeria species were used in this study. A description of the origins of the original isolates from which each line was derived is provided by El Sherry et al. (2015). The resulting lines used were as follows: 1) Eimeria adenoeides Guelph strain (see El-Sherry et al., 2014 for biological features of the line); 2) Eimeria dispersa Briston strain; 3) E. meleagrimitis USMN08-01 strain (see Long and Millard, 1979 and El-Sherry et al., 2014 for biological features); 4) E. meleagridis
USAR97-01 strain (see Matsler and Chapman, 2006 for biological features); 5) *E. gallopavonis* Weybridge strain (Hein, 1969 for biological features); and, 6) *E. gallopavonis* USKS06-01 strain. All lines were derived from parent isolates using the method of Remmler and McGregor, (1964) with the modification that agar plugs carrying a single oocyst were given within gelatin capsules orally to specific-parasite free poult. All animal experimentation was conducted in SPF birds at the Campus Animal Facility, (University of Guelph, Guelph ON, Canada); all experimental procedures were reviewed and approved by the University of Guelph’s Animal Care Committee and complied with the Canadian Council on Animal Care’s Guide to the Care and Use of Experimental Animals (2nd edition).

7.3.2. DNA EXTRACTION AND LONG PCR AMPLIFICATION

Purification of oocysts and genomic DNA extraction was carried out as previously described by Ogedengbe et al. (2011; 2013). Mitochondrial whole genome amplification for all five *Eimeria* species was initiated using two sets of specific primers that generated overlapping PCR fragments: 1) Cocci_MT-WG-F (5'-TACACCTAGCCAACACGAT-3') and Cocci_MT-WG-R (5'-GCAGCTGTAGATGGATGCTT-3'); and, 2) Inv_COI_262R (5'-AAWCGGCATCRTAGAATTG-3') and Inv_COI_461F (5'-CTAGCYATGGGATGTATTACTG-3'). Primers were designed from highly conserved regions within publically available mitochondrial genome sequences for *Eimeria* species infecting chickens (see Ogedengbe et al., 2013 for the species used in the primer design). The primer pairs Inv_COI_461F and Inv_COI_262R annealed 148 bp apart at bp 2069 2090 and bp 1920 1901 respectively, and the primer pairs Cocci_MT-WG-F and Cocci_MT-WG-R annealed 97 bp apart at bp 6322-6340 and bp 6224 6205, respectively, on the published mitochondrial genome sequence of *Eimeria mitis* [GenBank: KF501573]. Each pair of primers was used independently in a 50µl reaction. PCR reactions using QIAGEN LongRange PCR kit (QIAGEN, Valencia, CA, USA) protocol according to the manufacturer’s instructions with the modification that an additional 1.5mM MgCl2 was added to the PCR buffer provided by the manufacturer. For each *Eimeria* species, long PCR reactions consisted of ~200ng genomic DNA template (when using the Inv_COI_461F/Inv_COI_262R primers), and 25ng genomic DNA template (when using primers Cocci_MT-WG-F and

67
Cocci_MT-WG-R), 1× LongRange PCR buffer, 4mM MgCl2, 500µM of each dNTP, 2U LongRange PCR enzyme mix and 0.4µM of each primer. The PCR reaction profile consisted of denaturing at 93°C for 3min followed by 35 cycles of 93°C for 15s, 50°C for 30s, 68°C for 6min with a final extension cycle of 68°C for 10min in an MJ mini thermal cycler (Bio Rad, CA, USA). PCR products were electrophoresed at 50V through a 0.8% agarose gel prepared with 1xTAE buffer containing ethidium bromide. DNA bands were viewed using UV transillumination (Spectronics Corporation, New York, USA) and their sizes were compared to a 100 bp to 10 kb DNA ladder (Bio Basic Inc., Mississauga ON, Canada). DNA bands were excised from the gel and purified using a QIA quick gel extraction and purification kit (Qiagen, Toronto ON, Canada) according to the manufacturer’s instruction.

7.3.3. SEQUENCING

Purified PCR products were sequenced in both directions using a primer-walking strategy to generate near-complete mitochondrial genomes essentially as described by Ogedengbe et al. (2013). Sequencing was carried out using the ABI PRISM 7000 Sequence Detection System (Applied Biosystem Inc., Foster City, CA, USA) at the Laboratory services Division, University of Guelph (Guelph, ON, Canada).

7.3.4. SEQUENCE DATA ASSEMBLY AND ANALYSIS

The de novo sequence assembler within Geneious bioinformatics software (Version 6.1 and later versions, available from http://www.geneious.com) was used to trim and assemble Sanger sequencing chromatograms into high quality contigs for the primary PCR product from each species. To complete each mt genome, PCR products were generated using a reverse primer downstream of the original forward primer (i.e. Cocci_MT-WG-F or Inv_COI_461F) and a forward primer upstream of the original reverse primer (i.e. Cocci_MT-WG-R or Inv_COI_262R, respectively); primers were designed such that a minimum of 100bp of the resulting fragment overlapped the original long PCR product at each end. Each resulting PCR product was sequenced in both directions and the resulting consensus sequence was used to fill in the region between the two original long PCR amplification primers. The coding genes and rDNA fragments were first mapped by comparison with other Eimeria mt genomes (i.e. E.
Additional putative rDNA fragments were identified by comparing well conserved unannotated regions found in all of the aligned *Eimeria* sp. genomes to the mt genome of *Plasmodium falciparum* (M76611). Sequence identity between such conserved regions and rDNA fragments from *P. falciparum* greater than 60% were mapped as putative rDNA fragments. Putative start and stop codon positions for each of the coding DNA sequences (CDS) were identified following methods previously described by Ogedengbe et al. (2013). Translations using the mold/protozoan mitochondrial codon translation (i.e. translation_table_4) were searched using Blastp against the non redundant sequence database to confirm the identity of the translation product produced by each CDS. Base compositions and nucleotide changes within the CDS among the six mt genome sequences were analysed from within the Geneious software package.

### 7.3.5. Phylogenetic Analyses

The six newly generated, PCR–based mt genome sequences of *Eimeria* spp infecting turkey: *Eimeria dispersa* Briston strain; *E. meleagrimitis* USMN08 01 strain; *E. meleagridis* USAR97-01 strain; *E. adenoeides* Guelph strain and *E. gallopavonis* Weybridge strain; *E. gallopavonis* USKS06-01 strain were aligned with the 10 publically available complete mt genome sequences from seven *Eimeria* spp. infecting chickens and *Eimeria magna* that infects rabbits (i.e. all available apicomplexan taxa that had the same genome structure). Three sequences of *Eimeria mitis* (KCE409029; KC409030 and KC409031) generated from clones were not included in the phylogenetic analysis because of the likelihood of PCR artifacts in these sequences as documented by Ogedengbe et al. (2013). GenBank sequence accession numbers are indicated on the trees.

To permit whole genome alignments, all mt genome sequences were linearized at the same position, 85-87 nt upstream of the small subunit rDNA fragment SSU/A corresponding to the binding site of the Cocci_MT WG F primer. Linearized sequences were aligned based on the primary structure using the multiple sequence alignment algorithm implemented from within Geneious 6.1; indels downstream of the Cocci_MT
WG R primer binding site made unambiguous alignment in that region unlikely so the short sequences downstream of this primer binding region (44 to 97 bp depending on the *Eimeria* sp.) were not included in subsequent phylogenetic analyses using whole genome sequences.

Regions between rDNA fragments contained frequent indels that made unambiguous alignment of these regions difficult and the CDS for the three genes contained the majority of the genetic diversity found within the mt genomes. For these reasons, we chose to use concatenated CDS for CytB, COI and COIII (or their corresponding amino acid sequences) as datasets for phylogenetic analyses. The sequence data was thus partitioned into 3 datasets as follow: 1) a global nucleotide sequence data set for all 16 whole genome sequences (after removal of the short regions downstream of the Cocci_Mt WG R primer); 2) concatenated DNA sequences for the 3 CDS; 3) concatenated amino acid (aa) translations of the 3 CDS.

Phylogenetic analyses were performed on all 3 data sets using three tree building methods, Bayesian analysis (BI) (Huelsenbeck and Ronquist, 2001) performed using MrBayes (Version 3.2.), Maximum Parsimony (MP) using PAUP 4.0 (Swofford, 2003) and Maximum Likelihood (ML) using PhyML (Guindon et al., 2010) executed from within the Geneious bioinformatics software package (Version 6.1 and later versions).

Data set one, consisting of whole mt genome nucleotide sequences (excluding the short regions downstream of the Cocci_MT WG R primer), for all 16 whole genome sequences and data set two, consisting of concatenated CDS were analysed using all three tree-building methods (BI, ML or MP). Selection of the best fit evolutionary model for the BI and ML analyses was evaluated in both MrModeltest v2.3 (Nylander J. A. A. 2004. MrModeltest v2. Program; distributed by the author, Evolutionary Biology Centre, Uppsala University) and MEGA (Tamura et al., 2011). For the Bayesian analyses, Markov Chain Monte Carlo was performed for 1,000,000 generations with four chains and heated chain temperature of 0.2. The burn-in length was set at 400,000 and subsample frequency of 1000 (Huelsenbeck and Ronquist, 2011; Ronquist and Huelsenbeck, 2003). For the ML analyses, 500 bootstrap replicates were calculated to
estimate node support. In the MP analyses, characters were unordered and given equal weight; trees were searched using the branch and bound search algorithm.

Data set three, consisting of concatenated amino acid (aa) translations of the 3 CDS was analysed with the same three tree-building methods. The empirical Jones-Taylor-Thornton (JTT) model of amino acid substitution with gamma distribution frequency (G+F) for all sites by Jones et al. (1992) was selected for the ML and BI analyses. Substitution models were assessed in MEGA (Tamura et al., 2011). Where outgroup rooting was appropriate, the *Eimeria magna* mitochondrial genome sequence [GenBank: KF419217] was used as the functional outgroup.

Although sequences were truncated to remove the short indel-rich region downstream of the reverse primer for the phylogenetic analyses, complete genome sequence alignments were used within Geneious for calculating the pairwise genetic distances and number of nucleotide differences among the six newly sequenced genome sequences.

7.4. RESULTS

7.4.1. SIX MT GENOMES FROM FIVE *EIMERIA* SPECIES INFECTING TURKEYS

The six complete mitochondrial genome sequences obtained from direct sequencing of PCR products from five *Eimeria* spp infecting turkeys varied modestly in their lengths: *Eimeria adenoeides* Guelph strain [GenBank:KJ608415, 6211bp]; *Eimeria dispersa* Briston strain [GenBank:KJ608416, 6238bp]; *Eimeria meleagridis* USAR97-01 strain [GenBank:KJ608418, 6212bp]; *Eimeria meleagrimitis* USMN08 01 strain [GenBank:KJ608414, 6165bp]; *Eimeria gallopavonis* Weybridge strain [GenBank:KJ608413, 6215bp ]; and, *Eimeria gallopavonis* USKS06-01 strain [GenBank:KJ608417, 6215bp]. Base composition was A/T biased in all species (Table 7.1). Pairwise comparisons of the 6 aligned mt genome sequences from *Eimeria* spp. infecting turkeys (Table 7.2) indicate a high degree of sequence identity among these new genome sequences; 5311 nucleotide positions (84.9% of the aligned sequence lengths) were invariant among all 6 genome sequences. There was no intraspecific variation noted between two strains of *E. gallopavonis* (Weybridge and USKS06-01) whose complete mt genomes were identical. The two closely related *Eimeria* spp.
causing ‘cecal coccidiosis’ in turkeys (i.e. *E. adenoeides* and *E. meleagridis*) demonstrated a genetic distance of 1.8% from each other and each was 3.1% divergent from the two strains of *E. gallopavonis* (Table 7.2). The physical form of the mitochondrial genomes was not directly assessed in this study, however, the mt genomes of these *Eimeria* species must be either linear concatenated or circular to permit successful PCR amplification of near full length mt genomes.

**Table 7.1.** Base composition of whole mitochondrial genomes from five *Eimeria* species infecting turkeys.

<table>
<thead>
<tr>
<th>Parasite</th>
<th>A/T (%)</th>
<th>Adenine (%)</th>
<th>Cytosine (%)</th>
<th>Guanine (%)</th>
<th>Thymine (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Eimeria adenoeides</em> Guelph</td>
<td>64.6</td>
<td>29.8</td>
<td>18.4</td>
<td>17.0</td>
<td>34.8</td>
</tr>
<tr>
<td><em>Eimeria dispersa</em> Briston</td>
<td>65.5</td>
<td>29.9</td>
<td>17.6</td>
<td>16.8</td>
<td>35.6</td>
</tr>
<tr>
<td><em>Eimeria gallopavonis</em> USKS06-01</td>
<td>64.9</td>
<td>30.1</td>
<td>18.3</td>
<td>16.9</td>
<td>34.8</td>
</tr>
<tr>
<td><em>Eimeria gallopavonis</em> Weybridge</td>
<td>64.9</td>
<td>30.1</td>
<td>18.3</td>
<td>16.9</td>
<td>34.8</td>
</tr>
<tr>
<td><em>Eimeria meleagrimitis</em> USMN08-01</td>
<td>63.5</td>
<td>30.0</td>
<td>19.2</td>
<td>17.3</td>
<td>36.5</td>
</tr>
</tbody>
</table>

**Table 7.2.** Percent pairwise sequence identities (lower values) and number of nucleotide differences (upper values) for six mitochondrial genomes from five *Eimeria* species that infect turkeys.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Eimeria dispersa</em> Briston</td>
<td></td>
<td>742</td>
<td>472</td>
<td>486</td>
<td>486</td>
<td>486</td>
</tr>
<tr>
<td><em>Eimeria meleagrimitis</em> USMN08-01</td>
<td></td>
<td>88.1%</td>
<td></td>
<td>608</td>
<td>597</td>
<td>602</td>
</tr>
<tr>
<td><em>Eimeria adenoeides</em> Guelph</td>
<td>92.4%</td>
<td>90.2%</td>
<td></td>
<td>113</td>
<td>194</td>
<td>194</td>
</tr>
<tr>
<td><em>Eimeria meleagrimitis</em> USAR97-01</td>
<td>92.2%</td>
<td>90.4%</td>
<td>98.2%</td>
<td>191</td>
<td>191</td>
<td>191</td>
</tr>
<tr>
<td><em>Eimeria gallopavonis</em> Weybridge</td>
<td>92.2%</td>
<td>90.3%</td>
<td>96.9%</td>
<td>96.9%</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td><em>Eimeria gallopavonis</em> USKS06-01</td>
<td>92.2%</td>
<td>90.3%</td>
<td>96.9%</td>
<td>96.9%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>
7.4.2. Genome Organization

Genome content and organisation of all six *Eimeria* sp. mt genomes consisted of three protein coding genes (COI, COIII and CytB) interspersed with 15 LSU and 11 SSU rDNA fragments (Figure 7.1). Pairwise sequence alignments between individual rDNA regions annotated in the *Eimeria* sp. genomes and the corresponding rDNA fragments of *P. falciparum* (M76611) identified by Feagin et al. (Feagin et al., 2012) demonstrated pairwise sequence identities that ranged from 68.5% to 93.8%.

![Figure 7.1](image)

**Figure 7.1.** Mitochondrial genome organization of five *Eimeria* species (Apicomplexa; Eimeriidae) infecting turkeys, *Meleagris gallopavo* (Aves; Galliformes). Transcriptional direction and order of the three coding regions for CytB, COI and COIII (direction indicated by arrowed end) are identical to those of other *Eimeria* species. Fragments of LSU rDNA (lower LSU labels), SSU rDNA fragments (upper SSU labels) and unassigned RNA7 (lower label) were found in regions outside of the protein-coding genes. Naming of ribosomal DNA fragments follows the convention of Feagin et al. (2012, Table 1).

Searching conserved regions along the aligned mt genomes in the present study against the annotated *P. falciparum* mt genome identified three additional regions that are putative rDNA. The first two regions had high sequence identity to a single rDNA of *P. falciparum* encoding RNA14 (SSU/1) that appears to have been further fragmented on the *Eimeria* sp. mt genomes; the two resulting smaller fragments were found to map to two widely separated regions on these genomes. The first 29bp of RNA14 from the *P. falciparum* mt genome (bp 5576-5548 of M76611) has high pairwise sequence identity (~79%) to a region designated RNA14a on the mt genomes of all *Eimeria* spp. The following 41bp of RNA14 from the *P. falciparum* mt genome (bp 5547–5508 of M76611) has high pairwise sequence identity (75.6%) to a region designated RNA14b on the six mt genomes reported in the present study. The newly annotated rDNA
fragment RNA14a (SSU/1a) was found in reverse orientation starting at bp 5141-5110 (varies in each *Eimeria* sp. mt genome) and RNA14b (SSU/1b) was found in forward orientation starting at bp 3104–3111 (varies in each *Eimeria* sp. mt genome). The remaining conserved region for which high sequence identity was discovered with the mt genome of *P. falciparum* corresponded to RNA5 (SSU/9) annotated by Feagin et al. (2012). This putative rDNA fragment was found in reverse orientation starting at bp 6130-6199 (varies in each *Eimeria* sp. mt genome) and corresponded to bp 4724-4802 on the *P. falciparum* mt genome. Although the pairwise sequence identity between the complete RNA5 (SSU/9) regions on the *Eimeria* spp. and *P. falciparum* mt genomes was only 63.2%, both the 5’ and 3’ ends of these regions were highly conserved (i.e. 80%-85% sequence identity in the 20bp at each end of the region).

The COIII CDS was most divergent (76.3% identical sites across the six mt genome sequences). The COI and CytB CDS showed 81.2% and 81.8% identical sites, respectively. Of the 272 sites demonstrating variation among the 6 COI CDS examined, 239 were synonymous (KS) changes and 33 were non synonymous (KA) changes. The COIII CDS had 179 sites with variation (74 KS and 40 KA) and the CytB CDS had 197 variable sites (167 KS and 30 KA changes). The CDS were more divergent than the rDNA fragments (80.2% sequence identity over the 3279bp of the genomes identified as CDS versus 95.9% sequence identity over the 1880bp identified as rDNA regions). Nucleotide differences and indels were observed within some fragmented rDNA regions but were most commonly observed within intergenic regions (i.e. between regions annotated as CDS or rDNA).

7.4.3. START CODON DETERMINATIONS FOR COI, COIII AND CYTB

Start codon assignments were made by comparison with 13 publically available complete mt genomes from various *Eimeria* species and subsequent confirmation of appropriate open reading frames. In the six mt genomes obtained in this study, an ATG start codon for the CytB CDS beginning 214 or 215 bp downstream of the start of the Cocci_MT WG F primer binding site was shared among all *Eimeria* spp. for which complete mitochondrial sequences have been obtained. Preceding the ATG start codon was a poly–T-rich ‘GTTTATGTTTA’ motif that was conserved in all *Eimeria* spp. of
turkeys with the exception of *Eimeria meleagrimitis* USMN08-01. The latter sequence had a single substitution of ‘T’ with a ‘C’ producing a slightly different motif ‘GTTTATGTTCA’. A single stop codon, TAA, terminated the CDS for CytB, COI and COIII in all six mt genome sequences. Potential start codons for the COI CDS identified upstream of the highly conserved ‘Asn–His–Lys’ motif associated with the start of the heme–copper oxidase subunit I core region of COI were numerous for most of the new genome sequences except for *E. meleagrimitis* that had only 2 ORF’s that start upstream of that functionally conserved region (c.f. Ogedengbe et al., 2013). Of these two potential start codons for COI in the *E. meleagrimitis* sequence, only one potential start codon was shared among all *Eimeria* species; this ATD (ATG or ATA or ATT) start codon is located 27 bp upstream of the ‘Asn–His–Lys’ site in all *Eimeria* species sequenced to date. The start codon for the COIII CDS was determined to be a TTA codon located 14-20bp downstream of the LSU/1 (LSUA) region. Use of this conserved start codon produces a COIII product of 252 aa. In all *Eimeria* species studied thus far there is a poly-A- and poly-T-rich region located upstream of both the COI and COIII start codons.

7.4.4. PHYLOGENETIC ANALYSES

After trimming the alignment of whole genome sequences to remove the short indel-rich region downstream of the Cocci_WG MT R primer, the alignment of 16 available mt whole genome sequences used for phylogenetic analyses was 6416 bp in length, including gaps. The general time reversible model with discrete Gamma (GTR+I+G) distribution of nucleotide substitution (Rodriguez et al., 1990) was determined to be optimal for the BI and ML analyses. Figures 7.2 and 7.3 illustrate the phylogenetic relationships based on Bayesian inference (BI) and Maximum likelihood (ML); and Maximum parsimony (MP) models, among the 10 publically available complete mt genome sequences from eimeriid coccidia and the six newly generated complete mt genome sequences from *Eimeria* spp. infecting turkeys. The *Eimeria magna* mt genome sequence was used as a functional out-group in all phylogenetic analyses. Phylogenetic trees generated from aligned concatenated CDS for COI, COIII and CytB are illustrated in Supplementary Figure S1 (Appendix 2) for BI and ML.
analyses and Supplementary Figure S2 (Appendix 2) for the MP analysis. Trees generated using concatenated amino acid translations of the CDS matched the trees based on the concatenated CDS dataset under the same phylogenetic inference model (data not shown).

In the BI and ML trees, for global complete mitochondrial nucleotide sequences and the concatenated CDS, all *Eimeria* species causing ‘cecal coccidiosis’ in turkeys (i.e. *E. meleagridis*, *E. gallopavonis* and *E. adenoeides*) formed a monophyletic clade that was the sister group to *E. meleagrimitis*; the latter species infects the intestinal tract of turkeys excluding the ceca. The *Eimeria* species causing ‘cecal coccidiosis’ in chickens (i.e. *E. tenella* and *E. necatrix*) formed a monophyletic clade that was the sister clade to these four *Eimeria* species infecting turkeys. In the MP trees based on the same DNA sequences (complete genome or concatenated CDS), *Eimeria meleagrimitis* was the sister taxon to a monophyletic clade consisting of species causing ‘cecal coccidiosis’ in chickens and turkeys. In none of the analyses did all *Eimeria* species infecting turkeys form a monophyletic group; in all phylogenetic analyses *E. dispersa* branched near the base of the tree and was the sister taxon to all other *Eimeria* species within the functional ingroup. The *Eimeria* spp infecting chickens, excluding *E. tenella* and *E. necatrix*, formed a monophyletic clade in all analyses and all datasets (DNA and AA-based); however, the branching order within this monophyletic clade varied among analyses. All of these parasites (i.e. *E. acervulina*, *E. brunetti*, *E. mitis*, *E. praecox* and *E. maxima*) infect the intestinal tract of chickens outside of the cecal pouches.
Figure 7.2. Bayesian inference and maximum likelihood phylogenetic reconstructions using mitochondrial genome sequences of 16 *Eimeria* species. The analyses included 5 species infecting turkeys and 7 species infecting chickens and used *Eimeria magna* (a parasite of rabbits) as the functional outgroup to root the tree. Node support is indicated for BI (posterior probability, first number) and for ML (% bootstrap, second number) for all nodes with greater than 0.5 posterior probability. Neither the *Eimeria* species infecting chickens nor the *Eimeria* species infecting turkeys formed monophyletic groups. Both the BI and ML analyses supported monophyly of the 5 *Eimeria* species of chickens that do not usually invade the cecal pouches but branching order among these parasites was poorly resolved in both.
Figure 7.3. Maximum parsimony phylogenetic reconstruction using mitochondrial genome sequences of 16 Eimeria species. The analyses included 5 species infecting turkeys and 7 species infecting chickens and used Eimeria magna (a parasite of rabbits) as the functional outgroup to root the tree. Tree length was 3147 steps with a consistency index of 0.63 based on 927 parsimony-informative characters; percentage bootstrap support (500 replicates) is indicated at each node. The MP tree differed from the BI/ML tree only in the placement of E. meleagrisritis basal to a collection of lower intestinal tract parasites of chickens and turkeys. The MP analysis supported monophyly of the 5 Eimeria species of chickens that do not usually invade the cecal pouches.
7.5. DISCUSSION

The six newly reported mt genome sequences obtained in this study varied modestly in genome lengths (6165–6238 bp) and were comparable to the lengths (6148–6408 bp) of the mt genomes of *Eimeria* spp infecting chickens (Hikosaka et al., 2011; Lin et al., 2011; Liu et al., 2012; Ogedengbe et al., 2013) and rabbits (Tian et al., 2013). Genome organization of all mt genome sequences is highly conserved among eimeriid coccidia; however, eimeriid mt genome organization differs markedly from that of other apicomplexan mt genomes (e.g. Feagin, 2000; Hikosaka et al., 2010; Gray et al., 2004). No sequence differences (100% sequence identity) were recorded between the two strains of *Eimeria gallopavonis* (i.e. Weybridge strain and USKS06-01 strain) analysed in this study despite being isolated from different geographical regions. The two *E. tenella* sequences isolated from two geographical areas (Japan and China) also did not differ in their sequences (Hikosaka et al., 2011 and Lin et al., 2011, respectively). In comparison, the two *E. mitis* isolates from the US and China showed sequence differences at 32 positions; perhaps the longer domestication of the chicken host has permitted greater genetic variation in its parasites compared to the domesticated turkey.

The number, direction and lengths of the three CDS were identical in all six mt genome sequences obtained in the present work. Although the COI, COIII and CytB CDS have been annotated inconsistently in the publically available mt genome sequences, alignment of all 16 complete mt genomes from 13 *Eimeria* species demonstrated conserved CDS using the start codons identified by Ogedengbe et al. (2013) for *E. mitis* USDA50. An assessment of the three CDS across all six genome sequences yielded large numbers of nucleotide substitutions scattered within each gene.

Fragmented rDNA (from 16 to 188 bp in length) annotated in the present study were more highly conserved than the CDS, possibly due to functional constraints in the former. A single rDNA fragment (SSUA (SSU/4)) was found upstream of the CytB and COI genes, fifteen rDNA fragments were located between the COI and COIII genes and the ten remaining rDNA fragments were found between the COIII and the end of the mt genome. In addition to rDNA fragments identified in *P. falciparum* that had been
previously annotated as putative homologs on *Eimeria* sp. mt genomes (see, Ogedengbe et al., 2013), three regions of each *Eimeria* sp. mt genome had high sequence identities with rDNA fragments encoding RNA14 (SSU/1) or RNA5 (SSU/9) in *P. falciparum* (see, Feagin et al., 2012). A nearly complete rDNA encoding RNA5 (SSU/9) was located near the 3′ end of each genome and includes the binding site for the Cocci_MT-WG-R primer. The remaining two regions had high sequence identities to two portions of the rDNA encoding RNA14 (SSU/1) in *P. falciparum* (see, Feagin et al., 2012). However, this rDNA fragment appeared to have been further fragmented on the *Eimeria* sp. mt genomes and the two resulting smaller fragments (29bp and 41bp) were found to map to two widely separated regions on these genomes that we annotated as RNA14a (SSU/1a) and RNA14b (SSU/1b), respectively, on all six mt genomes reported in the present study.

All putative ribosomal fragments (fragmented LSU and SSU rDNA) were highly conserved among all *Eimeria* spp (Hikosaka et al., 2011; Lin et al., 2011 Ogedengbe et al., 2013; Liu et al., 2012; Tian et al., 2013; present study]. These putative rDNA fragments showed high sequence identity (from 62% to 93.8% pairwise identity) to functionally annotated rDNA fragments of *P. falciparum* M76611 (Feagin et al., 2012). Occurrence of fragmented and incomplete rRNA genes is not an uncommon phenomenon in apicomplexan parasites; similar fragmented rRNA genes have been reported in all other apicomplexan mt genomes examined to date (e.g. Feagin, 2000; Hikosaka et al., 2010; Feagin et al., 2012; Hikosaka et al., 2011; Lin et al., 2011). Although three additional conserved regions were annotated as putative rDNA fragments in the present study, other highly conserved regions in the six genome sequences remain unannotated but these comparatively conserved regions may represent as yet uncharacterized rDNA fragments.

Phylogenetic analyses under Bayesian, Maximum likelihood and Maximum parsimony evolutionary models using complete mt sequences or concatenated sequences from the three CDS from each mt genome did not support the conclusion that all *Eimeria* species infecting turkeys evolved from a common ancestor. Instead, although many turkey coccidia apparently share a common ancestor, at least one, *E. dispersa* was
found branching as the sister taxon to all other *Eimeria* spp in the functional ingroup. It is possible that *E. dispersa* may not have evolved within turkeys but rather arrived in that host via a host switch from some other avian host. *Eimeria dispersa* has been shown to infect both Bobwhite quail (*Colinus virginianus*) and turkeys (Tyzzer, 1929) and perhaps other hosts as well (Yabsley, 2009). In addition, the mt genome sequences suggest that the cecal coccidia of chickens (*E. tenella* and *E. necatrix*) are distantly related to the other *Eimeria* of chickens and are more closely related to some of the *Eimeria* spp that infect turkeys; this had been previously suggested on biological (Barta et al., 1997) and molecular (Miska et al., 2010) grounds. Analyses of the mt genome sequence data support the suggestion that *Eimeria* spp in chickens represent two distinct ancestral colonisations of the intestine. In one, *E. tenella* and *E. necatrix*, that appear closely related to a number of coccidia infecting turkeys, invaded the ceca of chickens; the remaining five *Eimeria* spp. infecting chickens are closely related using nu 18S rDNA (Ogedengbe et al., 2013; Barta et al., 1997; Miska et al., 2010), , partial mt COI sequences (Ogedengbe et al., 2013; Miska et al., 2010) or complete mt genome sequences (current study) and all of these species colonize regions of the intestine excluding the cecum.

Complete mt genome sequences could easily differentiate closely related parasites. For example, the pairwise genetic distance of *E. adenoeides* and *E. meleagrisidis* of turkeys and *E. tenella* and *E. necatrix* of chickens was 98.2% and 98.4%, respectively. Interestingly, the COI partial sequences for *E. adenoeides* (KCH strain) and *E. adenoeides* (KR strain) (Poplstein and Vrba, 2011) are 100% identical to the COI CDS of *E. adenoeides* (Guelph strain) and *E. meleagridis* (USAR97 01 strain), respectively, suggesting that the KCH and KR strains of *E. adenoeides* of Poplstein and Vrba, (2011) are distinct species rather than strains of a single species.

7.6. CONCLUSIONS

The mt genomes of *Eimeria* species infecting turkeys are similar with respect to genome size, organisation, start codon positions and overall base composition with all other *Eimeria* species. Complete mitochondrial genome sequences possess sufficient sequence variability for differentiating *Eimeria* species infecting turkeys or chickens.
and, in the three cases where more than one complete mt genome is available from a single species (i.e. *E. mitis*, *E. tenella* and *E. gallopavonis*), the intraspecific variation between mt genomes was much smaller (0–0.5%) than the genetic distance between that species and the most closely related *Eimeria* species (1.6 –3.2%). Genetic variability is concentrated within the three CDS encoding COI, COIII and CytB. This makes these mt genes of *Eimeria* spp. suitable (either as individual genes or as concatenated sequences) for species delimitation studies and phylogenetic analyses without the confounding presence of paralogous genome copies encountered with nu rDNA sequences (e.g. Ogedengbe et al., 2013; Vrba et al., 2011). The nature of the mt genome sequences, and particularly the CDS regions, of *Eimeria* spp. make the mt genome highly suited for development of diagnostic assays as well as, potentially, genetic markers for molecular epidemiology and phylogenetics of coccidia.
8. A LINEAR MITOCHONDRIAL GENOME OF CYCLOSPORA CAYETANENSIS (EIMERIIDAE, EUCCOCIDIORIDA, COCCIDIASINA, APICOMPLEXA) SUGGESTS THE ANCESTRAL START POSITION WITHIN MITOCHONDRIAL GENOMES OF EIMERIID COCCIDIA.

(The contents of this chapter have been published as follows: Mosun E. Ogedengbe, Yvonne Qvarnstrom, Alexandre J. da Silva, Michael J. Arrowood and John R. Barta, 2015. A linear mitochondrial genome of Cyclospora cayetanensis (Eimeriidae, Eucoccidiorida, Coccidiasina, Apicomplexa suggests the ancestral start position within mitochondrial genomes of eimeriid coccidia. Int. J. Parasitol. (2015), http://dx.doi.org/10.1016/j.ijpara.2015.02.006)

8.1. ABSTRACT

The near complete mitochondrial (mt)genome for Cyclospora cayetanensis is 6184 bp in length with three protein-coding genes (COI, COIII, CytB) and numerous lsrDNA and ssrDNA fragments. Gene arrangements were conserved with other coccidia in the Eimeriidae, but the C. cayetanensis mt genome is not circular-mapping. Terminal transferase tailing and nested PCR completed the 5´-terminus of the genome starting with a 21bp A/T-only region that forms a potential stem-loop. Regions homologous to the C. cayetanensis mt genome 5´-terminus are found in all eimeriid mt genomes available and suggest this may be the ancestral start of eimeriid mt genomes.

8.2. INTRODUCTION

Cyclospora cayetanensis is a protistan disease agent of humans that has been responsible for waterborne and large scale foodborne outbreaks worldwide; this apicomplexan parasite is now recognized as an emerging intestinal pathogen of public health importance (Karanja et al., 2007). In developed countries, the protist has been incriminated in diarrheal illnesses linked to imported foods (fresh uncooked vegetables and soft skinned fruits) and has been associated with diarrhea acquired through increased travel to endemic tropical regions (Shields and Olson, 2003). Particularly in the U.S., the two most recent diarrheal outbreaks caused by C. cayetanensis were solely associated with the consumption of fresh produce and affected 631 and 304 persons in 2013 and 2014, respectively (http://www.cdc.gov/parasites/cyclosporiasis/outbreaks/index.html). Cyclospora cayetanensis was first reported in stools of individuals suffering from protracted intermittent watery non-bloody diarrhea (CDC Report, 1991). Ortega et al. (1993) first identified the agent as an apicomplexan protist on the basis of oocyst morphology following sporulation and later on intracellular developmental stages. The organism was concluded to belong to the coccidian genus Cyclospora Schneider 1881,
family Eimeriidae Minchin, 1903 based on morphology of the sporulated oocyst that contains two sporocysts, each possessing a Stieda body and containing two sporozoites.

A single apicomplexan mitochondrial (mt) genome copy is approximately 6 – 7 kb in length (Gray et al., 2004). A variety of mt genome forms have been described among these parasites including circular-mapping mitochondrial genomes (e.g. circular in haemosporinids (Wilson and Williamson, 1997; Feagin et al., 2012) or linear concatemers (multiple mt genome copies joined end to end) in coccidia (Hikosaka et al., 2011)) as well as linear genomes with terminal inverted telomeric repeats in piroplasms (Hikosaka et al., 2010, 2012). Regardless of their structural forms, apicomplexan mt genomes usually possess three protein coding genes encoding cytochrome c oxidase subunit I (COI cytochrome c oxidase subunit III (COIII) and cytochrome b (CytB) as well as fragmented ssrDNA and lsrDNA. The present study reports on the sequence and structure of the mt genome of Cyclospora cayetanensis.

8.3. MATERIALS AND METHODS

Stool samples that were positive for the presence of C. cayetanensis by UV fluorescence microscopy were selected for the molecular studies. The samples were collected and used in accordance with the CDC Institutional Review Board (IRB) protocol entitled “Use of Human Specimens for Laboratory Methods Research”. Three samples were used for the whole mt genome sequencing: two samples from different time points (2011 and 2013) from an endemic area of southeastern Asia, plus a sample collected during the 2013 outbreaks in the USA. Partial COI and COIII genes were sequenced from five additional samples: two samples from the same southeastern Asian location (2012 and 2013); and, three samples from two different outbreaks in the USA during 2013. DNA was extracted using the Universal Nucleic Acid Extraction (UNEX) method as described by Shields et al. (2013) with some adjustments. Approximately 0.5 ml of stool was added to a matrix E bead beating tube (MP Biomedicals, Santa Ana, CA, USA) together with 60 µl of proteinase K (QIAGEN, Valencia, CA, USA) and 600 µl of UNEX buffer (Phthisis Diagnostics, Charlottesville, VA, USA). The tube was incubated at 56 °C for 15 min to allow for proteinase K activity. The mixture was homogenized in a FastPrep-24 tissue and cell disruptor instrument (MP Biomedicals) at a speed of 6.0
m/s for 1 min. The sample was then centrifuged at maximum speed (>13,000 g) for 1 min to pellet the debris. The supernatant was collected and passed through a DNeasy mini spin binding silica column (QIAGEN). Following two wash cycles using ethanol-containing wash buffers, the DNA was eluted from the column in 80 µl of AE buffer (QIAGEN). The eluted filtrate was further purified by passing through a Zymo-Spin IV-HRC column (Zymo Research Corp., Irvine, CA, USA). DNA samples were confirmed positive for *C. cayetanensis* by real-time PCR as described by Verweij et al. (2003).

Initial attempts to amplify near-complete mt genomes using methods that had worked reliably with various *Eimeria* spp. (see Ogedengbe et al., 2013, 2014) failed repeatedly with *C. cayetanensis*. Thereafter, shorter regions of the mt genome were amplified with primers targeting conserved regions of other apicomplexan mitochondrial genomes. The resulting PCR products were purified, sequenced directly using internal sequencing primers as necessary, and readily assembled into a partial mt genome using the de novo assembler within Geneious (www.geneious.com). Repeated attempts were made to complete the mt genome by amplifying across the ‘gap’ (assuming a circular or linear concatenated genome) with three pairs of additional amplification primers (Table 8.1). These efforts failed repeatedly with *C. cayetanensis* but succeeded in the case of coccidia in the genera *Eimeria*, *Caryospora* and *Isospora* (e.g. Lin et al., 2011; Ogedengbe et al., 2013, 2014; 2015a; Ogedengbe and Barta, 2015), all of which possess circular-mapping mt genomes; *Eimeria tenella* was used as positive control for all such PCRs due to its linear concatenated genome.

Under the assumption, based on PCR results, that the mt genome of *C. cayetanensis* was linear, bulk cellular DNA was tailed using a terminal deoxynucleotidyl transferase (TdT) tailing method previously described by Hikosaka et al. (2012) with minor modifications. Briefly, sample DNA (75 ng) in 5 µl of nuclease-free water was denatured at 94°C for 5 min and then immediately used in a 25 µl of 3´-tailling reaction consisting of 0.0125 µmol of dCTP (Clonetech Laboratories, Inc. Mountain View, CA USA), 7 U of TdT enzyme (Clonetech) and 0.02% BSA (w/v) in 1× TdT buffer (Clonetech) for 30 min at 37°C. At the conclusion of the tailing reaction, the TdT enzyme was heat inactivated at 65°C for 10 min (see Hikosaka et al., 2012). The poly-C-
tailed genomic DNA was then used as template for a pair of nested PCRs. In the first PCR, 2 μl of the end-labeled genomic DNA (i.e. 6 ng) was used as template in a reaction mixture containing 1.25 units of Platinum® Taq Polymerase (Invitrogen, Carlsbad CA, USA), 1× PCR buffer (Invitrogen), 2.5 mM MgCl₂, 200 μM dNTPs and 0.4 μM each of the required primers. For the amplification of the 5´-end of the genome, an mt genome-specific primer ‘q_Eim_CytB_398R’ (5´-CCCCAGWARCTCATYTGACCCCA-3´) was used with a poly-G-containing anchor primer ‘Telo_F_polyG’ (5´-GGCCACCGTGACTAGTACGGGGGGGGGGGGGG-3´). For the amplification of the 3´-end of the genome, an mt genome-specific WG_MT_5416F (5´-GGTCCAGATAAGCGATCTCATG-3´) was paired with the same poly-G-containing anchor primer (‘Telo_F_polyG’). Cycling conditions were initial denaturation at 95°C for 2 min followed by 40 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 45 s and final extension step at 72°C for 7 min. Following the primary PCR, 1 μl from each of the first PCRs was used as template for amplification reactions of 5´- and 3´-ends. The 5´-end was amplified using amplification primer Telo_F (5´-GGCCACCGTGACTAGTAC-3´) with mt genome-specific WG_MT_63R (5´-CTGGTATGGATACCGCTAAC-3´) under the following cycle conditions: initial denaturation at 95°C for 2 min, 40 cycles of 94°C for 30 s, 56°C for 30 s, 72°C for 20 s, and a final extension step at 72°C for 7 min. Similarly, the 3´-end was amplified with the primer Telo_F combined with mt genome-specific WG_MT_5813R (5´-AGGTGCTCAGGGTCTTACCG-3´) under slightly different cycle conditions: initial denaturation at 95°C for 2 min, 40 cycles of 94°C for 30 s, 63°C for 30 s, 72°C for 40 s and final extension step at 72°C for 7 min. After electrophoresis, amplification products were excised, purified and submitted for sequencing using the appropriate Cyclospora-specific primer used in the final amplification. The sequence data resulting from the mt genome tailing and subsequent nested PCR products was mapped onto the assembled partial C. cayetanensis mt genome obtained through direct PCR amplification. To allow alignment with related Eimeria sp. mt genome sequences, complete mt genome sequences from GenBank (conventionally notated as circular) were linearized such that position 1 of each sequence was at the region homologous to the 5´-terminus identified for C. cayetanensis (i.e. 5´-TATTnnAA…). Start and stop codon positions for each of three long open reading frames were identified using the Geneious
Version 6.1 (and later versions) software package. Annotation of the protein-coding regions within the near-complete mt genome sequence of C. cayetanensis followed that of the mt genome sequence from Eimeria mitis USDA 50 (Ogedengbe et al., 2013). Fragmented lsrDNA and ssrDNA were annotated following Feagin et al. (2012) as modified by Ogedengbe et al. (2015a).

8.4. RESULTS AND DISCUSSION

There were no observed sequence variations in the mt genomes of the C. cayetanensis isolates obtained from human feces examined in the present study regardless of the origin of the sample. Amplification of the most divergent regions of eimeriid mt genomes (the COI and COIII loci) illustrated that the five tested human isolates of C. cayetanensis (i.e. two southeastern Asian samples from 2012 and 2013; three USA samples from 2012 and 2013) were identical over all of the 860 bp of sequence obtained from the COIII locus and all of the 761 bp of sequence obtained from the COI locus. Other regions of the mt genome for which sequence data from multiple isolates were available showed no variation among any of the human isolates examined in this study (data not shown). The mt genome of C. cayetanensis is an apparently linear molecule approximately 6220 bp in length (see GenBank KP658101 for a near-complete sequence). Five overlapping fragments obtained with five pairs of amplification primers (Table 8.1) were obtained (1749 bp, 809 bp, 1978 bp, 959 bp and 1209 bp).

The full sequence length for the C. cayetanensis mt genome, including the completed 5’-end (but excluding 35 - 50 bp of the 3’-end estimated to be missing by comparison with the mt genome sequences of various Eimeria spp.) was 6184 bp (including the 3’ primer); thus the estimated full length of the mt genome would be approximately 6220 bp. Base composition of the mt genome was A (30.4%), T (36%), C (16.7%), G (16.9%) and thus the GC content was 33.6%. There was an A/T-rich terminus to the 5’-end of the genome that appeared capable of forming a closed hairpin loop structure. The C. cayetanensis mt genome encodes three protein-coding genes (CytB 1080 bp (237 bp - 1316 bp), ATG start codon; COI 1449 bp (1351 bp - 2799 bp), TTA start codon and COIII 753 bp (4362 bp - 5114 bp), ATG start codon). These coding
regions in the mt genome of *C. cayetanensis* had high mean pairwise sequence identities (90.4% in CytB, 90.3% in COI and 87.1% in Table 8.1. Amplification primers used to obtain the mitochondrial genome of *Cyclospora cayetanensis* and confirm its linear nature.

<table>
<thead>
<tr>
<th>FRAGMENT</th>
<th>PRIMER ID</th>
<th>NUCLEOTIDE SEQUENCE (5’-3’)</th>
<th>SIZE in bp (position within mt genome)</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragment 1</td>
<td>coci MT-WG-F</td>
<td>TACACCTAGCCAACACGAT</td>
<td>1807 (23 - 1829)</td>
<td>Ogedengbe et al, 2014b</td>
</tr>
<tr>
<td></td>
<td>Cyclo_COI_473R</td>
<td>ATACCCGCAAGAGCTAAACC</td>
<td>2030 (2442 - 4471)</td>
<td>This Study</td>
</tr>
<tr>
<td>Fragment 2</td>
<td>qPCR400-F</td>
<td>GTACAGGTTTTGTTGGAC</td>
<td>803 (1723 - 2525)</td>
<td>Ogedengbe et al, 2013</td>
</tr>
<tr>
<td></td>
<td>COI_1202R</td>
<td>CCAKRAYHGACCAAGAGATA</td>
<td>961 (4191-5151)</td>
<td>Ogedengbe et al, 2014b</td>
</tr>
<tr>
<td>Fragment 3</td>
<td>Cyclo_COI-1085F</td>
<td>TCCCCTCTAGGATGTTGCTT</td>
<td>2030 (2442 - 4471)</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>Cyclo_COI_113R</td>
<td>TTCCACCTTTGCTCTACGTG</td>
<td>1607 (4578 - 6184)</td>
<td>This Study</td>
</tr>
<tr>
<td>Fragment 4</td>
<td>WG-MT_4140F</td>
<td>AGAAAACCTAAAAATCATCATGT</td>
<td>431 (23-453)</td>
<td>Ogedengbe et al, 2013</td>
</tr>
<tr>
<td></td>
<td>Eim_COI_799R</td>
<td>GCATCCATCTACAGTGCTGGCG</td>
<td>343 (23-453)</td>
<td>Ogedengbe et al, 2014b</td>
</tr>
<tr>
<td>Fragment 5</td>
<td>Cyclo_COI_219F</td>
<td>AGCTTTCTTGAGGTGCAAC</td>
<td>No Product</td>
<td>Ogedengbe et al, 2013</td>
</tr>
<tr>
<td></td>
<td>coci MT-WG-R</td>
<td>TACACCTAGCCAACACGAT</td>
<td>431 (23-453)</td>
<td>Ogedengbe et al, 2013</td>
</tr>
<tr>
<td></td>
<td>WG_MT_344R</td>
<td>TACACCTAGCCAACACGAT</td>
<td>No Product</td>
<td>Ogedengbe et al, 2013</td>
</tr>
<tr>
<td></td>
<td>Cocci_MT_WG_F (Positive Control)</td>
<td>GTACAGGTTTTGTTGGAC</td>
<td>803 (1723 - 2525)</td>
<td>Ogedengbe et al, 2013</td>
</tr>
<tr>
<td>PCR Confirmation of Linear genome</td>
<td>WG_MT_344R</td>
<td>GTAGGAATCTRAATTCCCAACC</td>
<td>No Product</td>
<td>Ogedengbe et al, 2013</td>
</tr>
<tr>
<td></td>
<td>WG_MT_6219F</td>
<td>GCATCCATCTACAGTGCGG</td>
<td>No Product</td>
<td>Ogedengbe et al, 2013</td>
</tr>
<tr>
<td></td>
<td>WG_MT_5416F</td>
<td>GTAGGAATCTRAATTCCCAACC</td>
<td>No Product</td>
<td>Ogedengbe et al, 2013</td>
</tr>
</tbody>
</table>

1Note: Positions reported are relative to the complete mt genome of *Cyclospora cayetanensis* (GenBank KP658101); complete 5’-end was generated using TdT tailing followed by nested PCR and sequencing (see 2.3) but 3’-terminus has not been similarly completed.

COIII) to the corresponding coding DNA sequence (CDS) found in 17 mt genomes of various *Eimeria* spp. Numerous regions were identified as fragmented rDNA that appeared to encode 19 lsrDNA and 14 ssrDNA fragments. The rDNA fragmentation was irregular, with fragments ranging in length from 17 to 192 bp with the majority (20) interspersed between COI and COIII genes; the remaining 13 were located either before the CytB CDS (three fragments) or following the COIII CDS (10 fragments). The putative lsrDNA and ssrDNA fragments were annotated on the mt genome following the convention of Feagin et al. (2012) as extended by Ogedengbe et al. (2015a). Overall, pairwise comparisons of the complete *C. cayetanensis* mt genome to all available complete coccidian mt genomes demonstrated that the mt genome of *C. cayetanensis*
was most similar (92.5% pairwise identity) to *Eimeria dispersa* which has been demonstrated to infect turkeys (*Meleagris gallopavo*) and Northern Bobwhite quail (*Colinus virginianus*).

**Figure 8.1.** A) Aligned 5´- and 3´-termini of the *Cyclospora cayetanensis* mitochondrial genome aligned with the homologous regions in the mt genomes of other eimeriid coccidia; all comparative genomes have been linearized so that the first nucleotide of each mt genome matches the first base (‘T’) in the linear *C. cayetanensis* mt genome. A highly conserved region at the 5´-end with a consensus sequence of “5´-TATTTWAADA…” was found in all genomes included in the alignment; all eimeriid genomes illustrated have the potential to form hairpins at the 5´-“end” of their circularly-mapped genomes. At the 3´-end of the eimeriid mt genome units in all coccidia other than *C. cayetanensis* (data not available) there was an imperfect mirror repeat sequence (i.e. …CTGTNHTTTTGTC -5´, conserved mirrored bases underlined). The role, if any, of this terminal mirror repeat motif in the replication of eimeriid mt genomes is not known. The annotations listed below the *Caryospora bigenetica* mt genome sequence (terminal features, Cocci_MT-WG-F and R primers, and RNA9 and RNA5 SSU rDNA fragments) apply to all sequences in the alignment. B) Map of the *Cyclospora cayetanensis* mitochondrial genome illustrating the locations, directions and lengths of the 3 protein-coding genes (CytB, COI and COIII). The A-T rich 5´ end of the genome found in panel A is located at the start of the genome map (bp 1); the remaining smaller annotations illustrate locations and directions of the fragmented ribosomal RNA coding regions (LSU – dark annotated regions; SSU – light annotated regions) interspersed amongst the 3 protein-coding genes.

The *C. cayetanensis* mt genome appears to exist as a linear monomeric genome. Terminal transferase poly-C tailing followed by nested PCR and sequencing demonstrated a non-repetitive, 21 bp long region consisting of only A and T bases at the 5´-end of the *C. cayetanensis* mt genome that has the potential to form a tight hairpin loop (i.e. 5´-TATTTTTAATATAAAAAATTTA…; complementary regions underlined). Alignment of this region with the homologous regions from the mt genomes of coccidia belonging to several genera (Figure 8.1) illustrate that all of these parasites have
A/T-rich regions with many having a similar ability to form hairpin structures with a single terminal-free base. The sequence and structure of the 3′-terminus of the mt genome of *C. cayetanensis* could not be determined using the same methods used to obtain the 5′-end of the genome despite repeated attempts. When the sequence data at the presumed 3′ terminal end of other coccidia were examined, we observed a conserved, imperfect palindromic sequence motif ′CTGTTATTTTGTC′ at the 3′-end of all of these linearized mt genomes (Figure 8.1).

The mitochondrial genome of *C. cayetanensis* was highly conserved with respect to genome length, CDS content, transcriptional directions and lsrDNA and ssrDNA fragments compared with the mt genomes of *Eimeria* spp. infecting various hosts (e.g. Lin et al., 2011; Ogedengbe et al., 2013, 2014), an *Isospora* sp. infecting canaries (see Ogedengbe et al., 2015a) or *Caryospora bigenetica* (see Ogedengbe and Barta, 2015) that infects viperid snakes and various mammals. The mt rDNA fragments of *C. cayetanensis* had pairwise sequence identities of 68.2% to 97.6% compared with homologous rDNA fragments of *P. falciparum* (M76611) annotated functionally by Feagin et al. (2012). Fragmented rRNA genes are a common feature of apicomplexan mt genomes (Hikosaka et al., 2010; Feagin et al., 2012); the reason for the rDNA fragmentation is unknown. Similar to all apicomplexan mt genomes sequenced thus far, transfer rRNA are absent in the *C. cayetanensis* mt genome. The start codon usage for COIII and CytB genes was conventional and corresponded to the start codons in *Eimeria* spp. (Ogedengbe et al., 2013, 2014). The start codon ATG for the COI gene was identified at a position (1327 bp) in other *Eimeria* spp which does not exist in *C. cayetanensis*. Instead there is a TTA start codon two codons upstream of the putative start codon site found in all mt genomes of *Eimeria, Caryospora* and *Isospora* spp. examined to date (e.g. Hikosaka et al., 2011; Lin et al., 2011; Ogedengbe et al., 2013, 2014, 2015a; Ogedengbe and Barta, 2015). There was a lack of sequence variation observed in a number of isolates of *C. cayetanensis* at the mt COI and COIII loci even when obtained from geographically distant locations. Such lack of intraspecific variation at these loci has been reported for distinct geographic isolates of some *Eimeria* spp. (e.g. *E. tenella* and *Eimeria gallopavonis* have no intraspecific variation over the entire mt genome length) whereas *E. mitis* demonstrates limited genetic variation at these genetic
loci. Based on alignments of complete, annotated mt genome sequences from many coccidia, any intraspecific variation among *C. cayetanensis* isolates is likely to be found in the intergenic regions rather than within the coding regions (for both proteins and rRNA fragments).

Initial attempts to amplify a near complete mt genome sequence of *C. cayetanensis* using outward-facing primers within the COI CDS failed repeatedly; this approach had been used successfully with numerous *Eimeria* spp. and other eimeriid coccidia. Consequently, starting with known partial COI and COIII sequences, primers in highly conserved regions of the mt genomes of various eimeriid coccidia were designed until a near-complete genome was obtained. Ultimately, five overlapping PCR fragments were generated in this study to obtain a near-complete mt genome for *C. cayetanensis*. The physical structure of the *C. cayetanensis* mt genome apparently followed neither circular nor linear concatemer genome topology (i.e. circular-mapping); these latter forms usually permit the contiguous sequencing to obtain a single mt genome copy either by running around the circular genome or from one copy to the next in linear concatenated forms. Instead, the mt genome of *C. cayetanensis* appears to be an ‘open circle’ (i.e. linear) and this linear form made the amplification across the open region impossible. Different PCR attempts with reverse primers with confirmed binding sites near the 5´-end of the genome paired with forward primers with confirmed binding sites near the 3´-end of the genome all failed. To test whether a dimer consisting of a concatenated forward and reverse copy of the genome existed, PCRs using single forward primers near the 3´-end of the known genome were attempted; none produced any amplification products (data not shown). Although we conclude that the mt genome of *C. cayetanensis* is a linear monomer, the number of copies of the mt genome in each parasite was not determined.

Unlike the mt genomes of other coccidia examined thus far, TdT tailing, nested PCR and sequencing were required to complete the 5´-end of the *C. cayetanensis* mt genome; the first 21 bp of the mt genome consists of only A and T bases that had the potential to form a hairpin loop with a single free terminal T. Alignment of the *C. cayetanensis* mt genome with a number of other coccidian mt genomes demonstrated
homologous regions in all mt genomes examined, including an invariant “TATTT” region shared among all mt genomes that matched the 5’-terminus of *C. cayetanensis* (see Figure 8.1).

In each mt genome, a sequence permitting a similar hairpin structure was observed with sequence heterogeneity within the fold region of the loop. Our conclusion is that this hairpin region identified at the 5’-end of the *C. cayetanensis* mt genome is the likely start of all eimeriid mt genomes, even when arranged in concatemers. In future, all circular-mapping mt genomes from eimeriid coccidia should use this well-conserved “TATTT” region as the 5’-end of each genome to establish a standardized ‘reading frame’ for these eimeriid mt genomes. Interestingly, a 45 bp stretch of the mt genome of *Plasmodium falciparum* (bases 51 - 95 of GenBank M76611) possesses a region capable of forming a hairpin followed by a region of high sequence identity with the Cocci_WG-MT_F primer (RNA9/SSU8), suggesting that such a convention might be able to be applied more broadly within the Apicomplexa. Kosa et al. (2006) suggested that A/T-rich stretches at the 5’- and 3’-ends of the linear mt genomes of a number of *Candida* spp. may mediate homologous recombination events that can generate circular or concatenated mt genomes from ancestral linear mt genomes. Such an A/T rich sequence was found at the 5’-terminus of the *C. cayetanensis* mt genome but the 3’-end remains to be determined; repeated attempts to tail this end of the genome using TdT proved fruitless. Physical mapping of the genome may be useful to confirm the structure of this genome in future.

Molecular analyses of nuclear ssrDNA sequences suggest that *C. cayetanensis* is closely related to other coccidia, especially members of the genus *Eimeria* (e.g. Relman et al., 1996; Shields and Olson, 2003) in the family Eimeriidae. Some authors have even suggested *C. cayetanensis* be reclassified as an *Eimeria* sp. (Pieniazek and Herwaldt, 1997), although this would necessitate a broadening of genus definition to include a range of oocyst morphologies. Although numerous *Cyclospora* spp. have been described, sequence data exist for only four (i.e. *C. cayetanensis*, *Cyclospora papionis*, *Cyclospora colobi* and *Cyclospora cercopithei*) and these sequences are restricted to the nuclear ssrDNA and internal transcribed spacer (ITS) loci (Eberhard et al., 1999;
Zhou et al., 2011). The potential presence of paralogous gene copies in the nuclear rDNA and ITS loci as well as high sequence divergence at the ITS locus (loci) makes these genetic targets less than ideal for delimiting closely related coccidian parasites (Adams et al., 2000; El-Sherry et al., 2013). In contrast, mt CDS sequences (e.g. partial or complete COI sequences) have been used successfully for delimiting closely related coccidia (Ogedengbe et al., 2011, 2014; El-Sherry et al., 2013). Nuclear ssrDNA sequences place the protozoan parasite *C. cayetanensis* in the apicomplexan family Eimeriidae among various *Eimeria* and *Isospora* spp. The overall similarity of the mt genome content and sequence of *C. cayetanensis* supports the view that *Cyclospora* spp. are closely related to other eimeriid coccidia such as *Eimeria* spp.

8.5. ACKNOWLEDGEMENTS

This work was supported by scholarship support to MEO from the Ontario Veterinary College, Canada and a Discovery Grant (DG 400566) from the Natural Sciences and Engineering Research Council of Canada (NSERC) to JRB. This work was supported by the Centers for Disease Control (CDC) Initiative for Advanced Molecular Detection and Response to Infectious Disease Outbreak, USA. The authors acknowledge Julia Whale for her technical support.
9. THE COMPLETE MITOCHONDRIAL GENOME OF CARYOSPORA BIGENETICA (EIMERIIDAE, EUCCOIDIORIDA, COCCIDIASINA, APICOMPLEXA)

(The contents of this chapter have been published as follows: Mosun E. Ogedengbe, and John R. Barta, 2015. The complete mitochondrial genome of Caryospora bigenetica (Eimeriidae, Eucoccidiorida, Coccidiasina, Apicomplexa. Mitochondria DNA. Early Online: 1–2 doi: 10.3109/19401736.2015.1015007)

9.1. ABSTRACT

The 6313bp complete mitochondrial (mt) genome of Caryospora bigenetica was sequenced directly from PCR products. The mt genome was comparable in size, gene content and order to those of other eimeriid coccidia (e.g. Isospora or Eimeria species). Three protein coding genes encoding COI, COIII and CytB were identified; numerous rDNA fragments (19 LSU and 14 SSU) were interspersed among the CDS. Nucleotide composition was A+T biased (66%). The mitochondrial genomes of eimeriid coccidia appear to share the same gene order and content; mt genome sequences can provide molecular data useful for diagnostics, taxonomy and phylogenetic relationships of eimeriid coccidia.

9.2. INTRODUCTION

The coccidium Caryospora bigenetica Wacha and Christiansen 1982 undergoes typical endogenous development within the intestinal tract of its viperid definitive hosts (e.g. Timber rattlesnakes [Crotalus horridus] or Massasauga rattlesnakes [Sistrurus catenatus]). Oocysts of C. bigenetica shed by snakes that sporulate to infectivity can infect mice and other mammals (e.g. pigs or dogs) through ingestion (Euzeby, 1991) resulting in a complete extraintestinal life cycle that produces encysted sporozoites (‘caryocysts’) in these secondary hosts. Clinical signs can be observed in some secondary hosts (Lindsay et al., 1988a) and zoonotic infections in humans through consumption of improperly processed meat may be possible (Lindsay et al., 1988b). Rattlesnakes and susceptible secondary hosts can become infected by ingesting caryocysts within the tissues of infected secondary hosts (e.g. rodents or canids) (Douglas et al., 1992).
9.3. MATERIALS AND METHODS

Oocysts of *C. bigenetica* were purified from the feces of a captive Massasauga rattlesnake using standard methods (Ryley et al., 1976). Sporulated oocysts (Figure 9.1a) possessing a single polar granule (Figure 9.1b) measured (10.9 [10.1 – 12.5] × 10.7 [9.4 – 12.1] μm, n=33) and sporocysts measured (9.6 [8.9 - 11.01] × 7.4 [6.8 - 8.2] μm, n=33); the oocyst dimensions and their possession of a single polar granule matched the species description for *C. bigenetica* by Wacha and Christiansen (1982). Following DNA isolation, the complete mitochondrial (mt) genome of *C. bigenetica* (KP658102, 6313bp) was obtained using mt-specific primers reported previously (Ogedengbe et al., 2014).

9.3. RESULTS AND DISCUSSION

Gene content and directions within the *C. bigenetica* mt genome (Figure 9.1c) were identical to other eimeriid coccidia available (Lin et al., 2011; Tian et al., 2013; Ogedengbe et al, 2013; 2014; 2015a). Like most Apicomplexa, only three protein coding genes were located within the mt genome: cytochrome c oxidase subunit I (COI); cytochrome c oxidase subunit III (COIII); and cytochrome B (CytB). The COI and CytB genes used ATG start codons and COIII used TTA; all protein coding genes used TAA termination codons as reported previously for many eimeriid coccidia (e.g. Ogedengbe et al., 2014). The mt genome also encoded 19 LSU and 14 SSU rRNA fragments. Nomenclature for fragmented rDNA followed Feagin et al. (2012) as updated by Ogedengbe et al. (2015, submitted). No transfer RNA was present. Genome content was A+T biased (66%) with 29.5% A, 17.1% C, 16.9% G and 36.5% T. Protein-coding genes formed 52% of the total length of the genome with rDNA fragments found interspersed among these genes.

Complete nuclear small ribosomal DNA (18S rDNA genes, see Barta et al., 2001) and partial COI gene sequences (Ogedengbe et al., unpublished data) place *C. bigenetica* among the eimeriid coccidia (e.g. species in the genera *Eimeria*, *Isospora*, *Cyclospora*, etc.), usually as the sister taxon to *Lankesterella minima* (Lankesterellidae), suggesting close evolutionary relationships among these coccidia. Mitochondrial DNA sequence data are highly suited for species delimitation and analysis of diversity among
apicomplexan taxa (Ogedengbe et al., 2011, Pacheco et al., 2013). The complete mitochondrial genome sequence of *Caryospora bigenetica* may provide the data necessary to clarify taxonomic and phylogenetic relationships of this parasite to other eimeriid coccidia within the Apicomplexa.

**Figure 9.1.** *Caryospora bigenetica* (Apicomplexa; Eimeriidae) from a Massasauga rattlesnake (*Sistrurus catenatus*): a) Sporulated oocyst with a single sporocyst possessing a Stieda body and containing 8 sporozoites surrounding a sporocyst residuum (scale bar = 10 μm); b) Sporulated oocysts have a single refractile granule (arrow, scale bar = 10 μm); c) Diagrammatic representation of the mitochondrial genome of *C. bigenetica* (KP658102, 6313bp). The genome possesses three protein-coding genes (CytB - Cytochrome B, COI - cytochrome *c* oxidase subunit I, and COIII - cytochrome *c* oxidase subunit III) and numerous large subunit (blue) and small subunit (red) rDNA fragments; genes encoded in the forward direction are mapped at the bottom and those encoded in the reverse direction are mapped at the top. A short stem-loop structure is located at the 5′-start of the genome (grey) and an imperfect mirror repeat is located at the 3′-terminus (grey).
10. THE COMPLETE MITOCHONDRIAL GENOME SEQUENCE OF AN ISOSPORA SP. (EIMERIIDAE, EUCCOCRIDIORIDA, COCCIDIASINA, APICOMPLEXA) CAUSING SYSTEMIC COCCIDIOSIS IN DOMESTIC CANARIES (SERINUS CANARIA LINN.)

(The contents of this chapter have been published as follows: Mosun E. Ogedengbe, Marina Brash and John R. Barta, 2015. The complete mitochondrial genome sequence of an Isospora sp. (Eimeriidae, Eucoccidiorida, Coccidiasina, Apicomplexa) causing systemic coccidiosis in domestic canaries (Serinus Canaria Linn. Mitochondria DNA. Early Online: 1–3 doi: 10.3109/19401736.2015.1018201)

10.1. ABSTRACT

We report a complete mitochondrial genome sequence for an Isospora sp. causing systemic coccidiosis in canaries, Serinus canaria. The A+T rich (65.2%) genome was 6216 bp in length and possessed 3 protein coding genes, (COI; COIII and CytB), 19 LSU and 14 SSU rDNA fragments, including 1 newly identified putative LSU fragment. Arrangement of coding regions was identical to that of available Eimeria sp. mt genomes and start codon usage for protein coding genes was conventional. The similar mitochondrial genome sequences and structures of Isospora and Eimeria species confirm the close relationship between these eimeriid genera of apicomplexan parasites.

10.2. INTRODUCTION

Isospora species are monoxenous, host-specific coccidia (Apicomplexa) parasitizing Passeriformes (e.g. canaries, finches, sparrows, grosbeaks and starlings) and other birds transmitted via the fecal-oral route (Box, 1981; Levine, 1982; Schrenzel et al., 2005). Asexual (merogony) and sexual (gametogony) development within intestinal epithelial cells can result in morbidity and mortality, particularly in captive birds (Quiroga et al., 2000; Schrenzel et al., 2005; Cushing et al., 2011). Enteritis in passeriform birds caused by Isospora spp. is similar to Eimeria infections in immunologically naïve poultry; subclinical infections are common (Page and Haddad, 1995). Some Isospora species move extraintestinally during merogony causing systemic coccidiosis (e.g. Cushing et al., 2011; Hafeez et al., 2014). These forms can be confused with Atoxoplasma sp. parasites that are transmitted via haematophagous arthropods (Upton et al., 2001). Isospora sp. oocysts are diplosporocystic, tetrasporozoic structures (i.e. two sporocysts containing four sporozoites each) with Stieda bodies (Barta et al., 2005) unlike the diplosporocystic, tetrasporozoic without Stieda bodies formed by parasites in the Sarcocystidae (e.g. Toxoplasma or Hammondia, see Lindsay et al.,
Some *Isospora* species possess morphologically indistinguishable oocysts that makes specific identification based on fecal floats impossible (Hafeez et al., 2014). Molecular data from avian *Isospora* species are scarce and are usually nuclear small subunit rDNA sequences (e.g. Carreno and Barta, 1999). Schrenzel et al. (2005) also obtained short mt COI sequences in one of the first such uses of mt sequences in molecular studies of these coccidia.

In the present study, the complete sequence for a mitochondrial (mt) genome copy of an *Isospora* sp. causing systemic coccidiosis in captive canaries, (*Serinus canaria* Linnaeus) is described.

10.3. MATERIALS AND METHODS

DNA was extracted from infected tissues (liver, spleen, lungs and intestine) and then amplified using long-range PCR with coccidia-specific primers (Cocci_MT-WG-F and Cocci_MT-WG-R; see Ogedengbe et al., 2014) to obtain a near-complete (6168 bp) mt genome; the remainder (48 bp) was PCR-amplified to generate a 619 bp product with considerable overlaps (422 149 bp) with the primary amplicon. Sequences obtained by primer-walking were assembled using *de novo* sequence assembly within Geneious software (Version 6.1 and later, from http://www.geneious.com).

10.4. RESULTS AND DISCUSSION

The 6216 bp genome of the *Isospora* sp. (GenBank: KP658103) had three protein coding genes (COI, COIII and CytB) with 19 LSU and 14 SSU rDNA fragments but no regions encoding tRNAs; the genome shared >90% sequence identity with numerous *Eimeria* spp. which have identical genes and gene orders in their mt genomes (Ogedengbe et al., 2014). The start codons and their positions were conventional for COI, CytB and COIII (Table 10.1) as documented previously for other coccidian mt genomes (Ogedengbe et al., 2013); all CDS had TAA termination codons. Our amplification of overlapping PCR products covering the complete mt genome suggests that the *Isospora* sp. mt genome is physically either circular (like *Plasmodium* spp. [see Feagin et al., 2012]) or a concatemer of genome copies; the latter form is probable because concatenated mt genomes are reported for related *Eimeria* spp. (see Hikosaka et al., 2010).
Table 1.1. Protein coding genes and fragmented rDNA’s found in the complete mitochondrial genome of an *Isospora* sp. (Eimeriidae, Apicomplexa) causing systemic coccidiosis in canaries (*Serinus canaria* Linn).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Strand</th>
<th>Location</th>
<th>Size (bp)</th>
<th>Start Codon</th>
<th>Stop Codon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Putative Start (Stem-loop structure)</td>
<td>None</td>
<td>1 – 27</td>
<td>27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA9; SSU/8</td>
<td>+</td>
<td>35 – 81</td>
<td>47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSUA; SSU/4</td>
<td>+</td>
<td>123 – 199</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA23t; SSU (Tentative)²</td>
<td>+</td>
<td>200 – 234</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CytB CDS</td>
<td>+</td>
<td>245 – 1324</td>
<td>1080</td>
<td>ATT</td>
<td>TAA</td>
</tr>
<tr>
<td>COI CDS</td>
<td>+</td>
<td>1357– 2799</td>
<td>1443</td>
<td>ATG</td>
<td>TAA</td>
</tr>
<tr>
<td>RNA20 (Partial); LSU (Tentative)²</td>
<td>+</td>
<td>2805– 2824</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSU; LSU/11</td>
<td>+</td>
<td>2850 – 2961</td>
<td>112</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSU; LSU/12</td>
<td>+</td>
<td>2962 – 3067</td>
<td>106</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSU (Tentative)³</td>
<td>-</td>
<td>3076 – 3100</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA14; SSU/1</td>
<td>+</td>
<td>3141 – 3180</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSUC; LSU/4</td>
<td>-</td>
<td>3193 – 3209</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA22; SSU (Tentative)²</td>
<td>+</td>
<td>3233 – 3262</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSUF; LSU/12</td>
<td>+</td>
<td>3345 – 3405</td>
<td>61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA10; LSU/13 (Partial)</td>
<td>+</td>
<td>3420 – 3479</td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA11; LSU/5</td>
<td>+</td>
<td>3503 – 3551</td>
<td>49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSUD; LSU/10</td>
<td>+</td>
<td>3560 – 3624</td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA17; LSU/3</td>
<td>+</td>
<td>3659 – 3698</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA15; SSU (Tentative)²</td>
<td>+</td>
<td>3705 – 3734</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA13; LSU/10</td>
<td>+</td>
<td>3742 – 3771</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA6; LSU/15</td>
<td>+</td>
<td>3778 – 3835</td>
<td>58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSUD; LSU/8</td>
<td>-</td>
<td>3859 – 3939</td>
<td>81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA16; LSU (Tentative)²</td>
<td>-</td>
<td>3948 – 3970</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA8; LSU/5</td>
<td>-</td>
<td>3985 – 4076</td>
<td>92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA2; LSU/2</td>
<td>+</td>
<td>4094 – 4159</td>
<td>66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSUA; LSU/1</td>
<td>-</td>
<td>4174 – 4354</td>
<td>181</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COIII CDS</td>
<td>+</td>
<td>4369 – 5124</td>
<td>756</td>
<td>TTA</td>
<td>TAA</td>
</tr>
<tr>
<td>RNA19; SSU/7</td>
<td>-</td>
<td>5134 – 5162</td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA1; LSU/6</td>
<td>+</td>
<td>5196 – 5283</td>
<td>88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSUB; LSU/3</td>
<td>-</td>
<td>5359 – 5385</td>
<td>27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA3; LSU/7</td>
<td>-</td>
<td>5421 – 5493</td>
<td>73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA18; LSU/14</td>
<td>-</td>
<td>5507 – 5526</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSUB; LSU/6</td>
<td>-</td>
<td>5551 – 5569</td>
<td>119</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA7; LSU (Tentative)²</td>
<td>-</td>
<td>5674 – 5735</td>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSUE; LSU/9</td>
<td>-</td>
<td>5763 – 5954</td>
<td>192</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSUE; SSU/11</td>
<td>-</td>
<td>6058 – 6090</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA5; SSU/9</td>
<td>-</td>
<td>6109 – 6211</td>
<td>103</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imperfect Mirror Repeat</td>
<td>None</td>
<td>6204 – 6216</td>
<td>13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹ The naming of genes encoding rDNA fragments follows the identifications presented in Table 1 of Feagin et al. (2012);

² Some rDNA fragments identified by Feagin et al. (2012) (i.e. RNA7, RNA15, RNA16, RNA20, RNA22 and RNA23t) were assigned tentatively to LSU or SSU by comparison with the reference genome for *Rickettsia rickettsii* str. Hauke - NC_016911.1;

³ An additional putative LSU rDNA fragment located at 3100 - 3076 mapped at 73.1% to the reference genome for *Rickettsia rickettsii* str. Hauke - NC_016911.1 but was not found on the *P. falciparum* mt genome (M76611.1).
11. MOLECULAR PHYLOGENETIC ANALYSES OF TISSUE COCCIDIA (SARCOCYSTIDAE; APICOMPLEXA) BASED ON NUCLEAR 18S rDNA AND MITOCHONDRIAL COI SEQUENCES CONFIRMS THE PARAPHYLHY OF THE GENUS HAMMONDIA.

11.1. ABSTRACT

Partial mitochondrial cytochrome c oxidase subunit I (COI) sequences were generated from nine Toxoplasma gondii strains (CTG, GTI, MAS, ME49, PTG TgCatBr5, TgCatBr64 TgCgCal, and TgToucan), Neospora caninum (Strain NC1), Hammondia hammondi (Strain H.H–20), Hammondia heydorni or H. cf triffittae (several clinical isolates from domestic dogs), Cystoisospora felis, C. suis, C. canis, C. rivolta C. cf. ohioensis, Caryospora bigenetica, and Eimeria sp. All nuclear 18S rDNA sequences were obtained from previous data from GenBank with the exception of newly generated 18S rDNA sequences for H. heydorni, H. hammondi, C. suis, C. canis, C. cf. ohioensis, C. bigenetica, Nephroisospora eptesici; Sarcocystis neurona, Sarcocystis rileyi, Isospora gryphoni, Lankesterella minima and an unnamed Isospora sp. (ex. Serinus canaria) causing extra intestinal coccidiosis in domestic canaries. Four pairs of group-specific primers were used to target an approximately 500-900bp fragment of the coding region of the mitochondrial COI gene. Phylogenetic analysis using Maximum Likelihood and Bayesian methods based on aligned nuclear (nu) 18S rDNA sequences supported the monophyly of the Sarcocystidae as well as its two subfamilies Toxoplasmatinae and Sarcocystinae. The Toxoplasmatinae held two major supported clades as deduced from 18S rDNA sequences; the first contained a monophyletic clade for all Cystoisospora spp. and the second was an unresolved polytomy of Toxoplasma, Neospora and Hammondia spp. Phylogenetic analyses using aligned COI sequences generated in this study with publically available sequences confirmed the monophyly of the Sarcocystidae. COI-based phylogenetic analyses also supported monophyletic clades for all genera within the Toxoplasmatinae for which more than one sequence was available with the exception of the paraphyletic genus, Hammondia. The parasites infecting dogs, H. heydorni and H. triffittae, formed a monophyletic clade and sister group to N. caninum. Hammondia hammondi formed a monophyletic clade with T. gondii in a second clade. These observations support the use of mt COI sequences as a better genetic marker for identifying and delimiting tissue coccidia than nu18S rDNA
sequences and call into question the stability of generic assignments of parasites within the Toxoplasmatinae.

11.2. INTRODUCTION

Coccidia in the family Sarcocystidae, the tissue coccidia, are important pathogens of many vertebrates including, in some cases, humans (Velmirugan and Dubey, 2008). Many species within the subfamily Toxoplasmatinae (e.g. *T. gondii, N. caninum, H. heydorni* and *H. hammondi*) exhibit facultatively or obligatory heteroxenous life cycles in which merogony (sometimes including endodyogeny) occurs in an intermediate host resulting in the formation of extra-intestinal asexual cystic stages containing bradyzoites. Gamogony occurs in the intestine of definitive hosts that release unsporulated oocysts into the environment. Exogenous sporulation produces disporocystic tetrazoic oocysts without Stieda bodies (Carreno et al., 1998; Mugridge et al., 1999). Coprological identification of *N. caninum, T. gondii, H. heydorni* and *H. hammondi* is difficult because their unsporulated oocysts are morphologically indistinguishable for all practical purposes (Frenkel and Dubey, 1975a; Lindsay et al., 1997; Dubey and Lindsay, 1996). Further, coprophagy may result in oocysts appearing as pseudo-parasites in the feces of atypical hosts (e.g. Schares et al., 2005). Only some *Cystoisospora* spp. (e.g. *C. canis* and *C. felis*) possess distinctly larger oocysts that can be diagnosed reliably using microscopy.

Sarcocystid coccidia all produce oocysts that contain 2 sporocysts, each containing 4 sporozoites (so-called ‘isosporoid coccidia’); the sporocysts lack Stieda bodies. Historically, all coccidia producing isosporoid oocysts were assigned to a single coccidian genus, *Isospora* Schneider 1881; however, isosporoid coccidia with Stieda bodies within their sporocysts have now been recognized to belong to the family Eimeriidae (see Carreno et al., 1998; Carreno and Barta, 1999; Barta et al., 2005). The taxa within the well-supported Sarcocystidae can be divided into two subfamilies based on whether or not asexual replication occurs in the definitive host prior to gametogony; members of the Toxoplasmatinae have merogony preceding gametogony whereas members of the Sarcocystinae do not (Frenkel, 1977). There are only two genera in the Sarcocystinae (*Sarcocystis* Lankester 1882 and *Frenkelia* Biocca 1968) distinguished by
the location of tissue cysts within intermediate hosts; the monophyly of these genera has been questioned repeatedly based on phylogenetic hypotheses erected using nuclear rDNA sequences (e.g. Mugridge et al., 2000; LSU rDNA, Morrison et al., 2004). Based on the ability to better resolve relationships among closely related coccidia using mt COI sequences (e.g. Ogedengbe et al., 2011; El-Sherry et al., 2013), the monophyly of the described genera in the Sarcocystinae might best be tested using this genetic locus.

The number of valid genera and species within the coccidian subfamily Toxoplasmatinae has been the subject of ongoing debate (e.g. Dubey et al., 1970; Overdulve, 1970a; 1970b; Frenkel, 1974; Frenkel and Dubey, (1975a, 1975b); Dubey, 1977; Frenkel, 1977; Mehlhorn and Frenkel, 1980; Dubey and Lindsay, 1996; Dubey et al., 1988; Blagburn et al., 1988; Carreno et al., 1998; Carreno and Barta, 1999; Dubey, 1999; Ellis et al., 1999; Tenter et al., 2000; Mehlhorn and Heydorn, 2000; Heydorn and Mehlhorn, 2001; Gjerde and Dahlgren, 2011; Schares et al., 2002; 2003; Dubey and Sreekumar, 2003; Abel, et al., 2006; Ajoka and Morrissette, 2009; Weiss and Dubey, 2009). Currently, there are 7 described genera recognized in the subfamily: Besnoitia Henry 1913; Cystoisospora Wenyon 1923, Frenkel 1977; Hammondia Frenkel 1974; Frenkel and Dubey 1975; Hyaloklossia Labbé 1896; Nephroisospora Wünschmann et al. 2010, Neospora Dubey et al. 1988; and, Toxoplasma Nicolle & Manceaux 1908.

Excluding the monotypic genera Nephroisospora and Hyaloklossia that shed sporulated oocysts, parasites in the remaining genera shed resilient, unsporulated oocysts in the feces of their definitive hosts. The assignment of the latter parasites to different genera depends on their interactions with their definitive and intermediate/paratenic hosts (see Frenkel, 1977).

Isosporoid coccidia that are facultatively heteroxenous and usually forms non-replicating monozoic cysts, (see Lindsay et al., 2014) are placed within the genus Cystoisospora. Monozoic cysts are infective only to the predatory definitive host; sporulated oocysts are infective to both paratenic (optional intermediate) and definitive hosts. Most isosporoid coccidia infecting mammalian definitive hosts that were previously assigned to the eimeriid genus Isospora (e.g. “Isospora felis” of cats or “Isospora ohioensis” of dogs) are actually Cystoisospora spp. (see Barta et al., 2005).

Species in the genera *Toxoplasma* and *Neospora* are facultatively heteroxenous isosporoid coccidia that use felids or canids, respectively, as definitive hosts. *Toxoplasma gondii* was described from tissue cysts (bradyzoites) observed in rodents in 1909. In addition to infections of definitive hosts by cysts (bradyzoites and tachyzoites), oocysts are also infective to both the definitive and intermediate hosts (Frenkel, 1977). The undescribed disease caused by *N. caninum* was first reported in dogs in Norway (Bjerkas et al., 1984). Dubey et al. (1988) proposed the name *N. caninum* to reflect its distinctiveness from *T. gondii*; Dubey et al. (2002a) summarized the biological, morphological, molecular and antigenic differences that distinguish these two parasites. Like the primarily monoxenous *Cystoisospora* spp., *N. caninum* and *T. gondii* can infect their definitive hosts via fecal-oral transmission of sporulated oocysts (see Dubey and Lindsay, 1996 and Dubey et al., 1970 respectively). Fecal-oral transmission in *N. caninum* was only experimentally induced and has not been confirmed as a valid mode of infection in naturally occurring transmission. Infections with both *N. caninum* and *T. gondii* rely heavily on heteroxenous transmission involving tachyzoite replication and bradyzoite development within cysts in a range of intermediate hosts (Dubey et al., 1970; Overdulve, 1970a; and Tenter et al., 2000).

The genus *Besnoitia* Henry 1913 was named for tissue coccidia that form thick-walled polyzoic cysts within connective tissue of their intermediate hosts. As an example, *Besnoitia wallacei* (Tadros and Laarman 1976) Frenkel 1977 was described from the oocysts in feces shed by a stray cat (definitive host) in Oahu, Hawaii (i.e. the WC-1170 isolate). *Besnoitia* spp. are obligately heteroxenous parasites. Within the intermediate hosts, tachyzoites appear in groups but their multiplication is limited (Wallace and Frenkel, 1975) before bradyzoite formation in fibroblasts. Infection of the definitive host is only by ingestion of cysts containing bradyzoites within the tissues of the intermediate hosts. In some species (i.e. *B. jellisoni* and *B. besnoiti*), limited
transmission from one intermediate host to the next is retained (Wallace and Frenkel, 1975; Frenkel, 1977).

Frenkel and Dubey, (1975a) erected the genus *Hammondia* to accommodate a feline coccidium, *H. hammondi*, with an obligatory two host life cycle. Cysts of *H. hammondi* containing infective bradyzoites form in skeletal and heart muscle of intermediate rodent hosts and these bradyzoites are infective only to the predatory definitive cat host. Once sporulated, oocysts shed by the definitive felid host are infective only to suitable intermediate hosts unlike oocysts of *T. gondii* that are infective to both the felid definitive and various intermediate hosts. Horizontal transmission of stages among intermediate hosts (as is typical of *T. gondii* and less commonly in *N. caninum* and some *Besnoitia* spp.) does not occur in *Hammondia* spp. as far as is known. *Hammondia heydorni* (Tadros and Laarman 1976; Dubey 1977), a parasite of canids, was characterized by Dubey, (1977) and Blagburn et al. (1988) from oocysts retrieved from feces of dogs. A second *Hammondia* sp. of canids, *H. trifittae*, has been recognized in European foxes by Gjerde and Dahlgren, (2011) who distinguished this parasite from other isosporoid coccidia of dogs (i.e. *H. heydorni* or *N. caninum*) by subtle differences in oocyst dimensions, genetic differences and the inability of *H. trifittae* to infect dogs. The separation of these two *Hammondia* spp. of canids was supported by later molecular comparisons (Dahlgren and Gjerde, 2010; Gjerde, 2011; Schares et al., 2002; 2003).

Most molecular phylogenetic studies of apicomplexan parasites in the family Sarcocystidae relied on small subunit (18S) rDNA sequences (Tenter and Johnson, 1997; Mugridge et al., 2000). The nuclear 18S rDNA locus is a comparatively useful genetic target from which to obtain sequence data due to its high copy number and conserved terminal sequences that are convenient for PCR primer design and amplification. The helical secondary structures of the small subunit are conserved in most organisms due to selection pressure and are particularly useful in the design of primers for intervening divergent regions (Morrison and Ellis, 1997; Mugridge et al., 2000; Ouvrard et al., 2000). These attributes and the relatively slow evolving nature of the 18S rDNA within the nuclear genome have made 18S rDNA an important and
widely exploited genetic target for biodiversity and species identification studies. In most eukaryotes, rDNA copies undergo concerted evolution that homogenizes copy to copy variability in the nuclear genome (Hillis, 1991). However, this is not always true for members of the Apicomplexa. Highly divergent paralogous copies of nuclear rDNA have been demonstrated within haemosporinid parasites (McCutchan et al., 1988, 1995; Li et al., 1997), piroplasms (Goethert et al., 2006), cryptosporidia (Le Blancq et al., 1997) and, more recently, eimeriid coccidia (Vrba et al., 2011; El-Sherry et al., 2013). The possibility of highly divergent rRNA genes (e.g. A, S, and O 18S rRNA genes that vary in both primary and secondary structures) could make determining homology at the gene level problematic.

Morrison et al. (2004) noted that complete 18S rDNA sequences were inadequate to resolve relationships among closely related tissue coccidia. Phylogenetic reconstructions based on 18S rDNA sequences support a monophyletic Sarcocystidae consistently (Ellis and Morrison, 1995; Carreno et al., 1998; Mugridge et al., 2000; Morrison et al., 2004); however, monophyly of individual genera within the Sarcocystidae is not supported in the same analyses. Unlike nuclear 18S rDNA sequences, comparatively short sequences obtained from the mitochondrial cytochrome c oxidase subunit I (COI) gene provide sufficient sequence divergence to clearly differentiate closely related coccidia. Ogedengbe et al. (2011) demonstrated that ~500-800bp partial mt COI sequences could better distinguish closely related coccidia (Eimeriidae) compared to complete or near-complete nuclear 18S rDNA sequences. Until recently (Gjerde, 2013a; 2013b), lack of primers capable of amplifying the mt COI locus from sarcocystid coccidia restricted the use of mt COI sequences to only eimeriid coccidia (e.g. Eimeriidae) and a few members of the Toxoplasmatinae (see Ogedengbe et al., 2011).

In this study we used existing and newly designed PCR primers to amplify and sequence portions of the mt COI gene from a variety of tissue coccidia (i.e. parasites belonging to the genera Toxoplasma, Neospora, Hammondia, Cystoisospora and Sarcocystis in the Sarcocystidae). Partial mt COI sequences, nuclear 18S rDNA sequences or a concatenation of sequences from both loci were utilized to infer
evolutionary relationships among these tissue coccidia and confirm the monophyly of some named genera in the family Sarcocystidae.

11.3. MATERIALS AND METHODS

11.3.1. PARASITE SOURCES

Parasites and parasite genomic DNA were obtained from a variety of sources. Fecal specimens containing oocysts of *C. felis*, *C. suis*, *C. cf. ohioensis*, *C. canis*, *C. rivolta* and *H. heydorni* were obtained from diagnostic fecal specimens submitted to the Animal Health Laboratory, Laboratory Services Division, University of Guelph (Guelph ON Canada). Oocysts in fecal specimens (for *C. felis*, *C. canis*, *C. suis*, *C. cf. ohioensis* and *C. rivolta*) from a variety of hosts were kindly provided by Dr. Donald Martin, (IDEXX Laboratories, Markham, Ontario, Canada); additional fecal samples containing *C. cf. ohioensis* were provided by Dr. Scott Weese (Department of Pathobiology, University of Guelph, Ontario). DNA samples of *Sarcocystis* spp. from experimentally infected hosts (i.e. *S. rileyi* 908131_Duck2 #4) or from tissue culture (i.e. *S. neurona* MIH2) were kindly provided by Dr Ben Rosenthal (USDA, Beltsville, MD, USA). Purified parasite genomic DNA was obtained for *H. hammondi* strain H.H-20 and *T. gondii* strains GT1, MAS, PTG, TgCat, CalBr64, CalBr5, TgCal and TgToucan prepared from scrapes of tachyzoite-infected cell cultures maintained by one of the authors in the Department of Microbiology, University of Tennessee; the origins of these lines have been described previously (Su et al., 2012). Purified parasite genomic DNA for *Cyclospora cayetanensis* was kindly provided by Dr Yvonne Qvarnstrom, United States Centers for Disease Control and Prevention, Atlanta, GA, USA.

11.3.2. IDENTIFICATION OF OOCYSTS.

Sporulated *Cystoisospora* oocysts collected and measured for this study were from dogs, cats, and pigs, identified as *C. canis*, *C. felis*, *C. rivolta* and *C. suis*. Small subspherical oocysts shed by dogs were considered part of the *C. cf. ohioensis* species complex and assigned the name *C. cf. ohioensis* to reflect the uncertainty in their species identification. In all cases prior to DNA extraction, initial assignment to species was on the basis of oocyst morphometrics. Oocysts images were captured using a Provis AX 70 photomicroscope (Olympus 95 Canada, Richmond Hill, ON) fitted with a digital
imaging device (Infinity3-1C, Lumenera Corporation Ottawa, ON, Canada) controlled using iSolution Lite image analysis software (Hoskin Scientific, Burlington, ON, Canada). Measurements are reported as means ± standard deviation in μm with the range in parenthesis.

11.3.3. OOCYST PURIFICATION AND DNA EXTRACTION

Oocysts were isolated from fecal samples using saturated salt (NaCl) flotation followed by a bleach treatment to remove exogenous DNA prior to parasite DNA extraction (Ogedengbe et al., 2013). DNA isolation was accomplished using glass bead disruption and DNAzol (Invitrogen, Carlsbad, CA) nucleic acid extraction as previously described (Ogedengbe et al., 2013). DNA concentration was estimated using Nanodrop 2000 spectrophotometer instrument (NanoDrop Products, Wilmington DE, USA) and stored at 4°C for immediate use or -20°C for later use.

11.3.4. MITOCHONDRIAL COI AND NUCLEAR 18S rDNA AMPLIFICATION PRIMERS

Initial PCR reactions attempted with ‘universal’ COI LCO1490 and HCO2198 primers that amplify many metazoan mt COI (Folmer et al., 1994) failed to amplify the COI locus from any species within the Toxoplasmatinae. Previously described coccidia-specific primers (Ogedengbe et al., 2011) were able to amplify some, but not all, members of the Sarcocystidae. Consequently, a pair of new Sarcocystidae-specific primers was designed using an alignment of all available COI sequences from tissue coccidia (data not shown). The COI sequence alignment was constructed using the ‘Translation Align’ algorithm in the Geneious bioinformatics software package (Version 6.1.8; http://www.geneious.com, Kearse et al., 2012). A pair of degenerate primers (Sdae- COI 260F, degeneracy 4; Sdae- COI 1147R, degeneracy 2) was designed from two comparatively conserved regions with the aid of Primer3 (Untergrasser et al., 2012) executed from within Geneious. The primers for 18S rDNA were Medlin A and Medlin B (Medlin et al., 1988), and ERIB1F forward and ERIB10R reverse primes as described by Barta et al. (1997). These previously published primers and newly designed primers were then used in various combinations to amplify ~500-900 bp fragments of the mt COI locus and ~1800bp fragments of the nuclear 18S rDNA (Table 11.1).
Table 11.1. Mitochondrial COI and nuclear 18S rDNA amplification primer combinations anneal temperatures (Ta) and expected product sizes for PCR of various tissue coccidia.

<table>
<thead>
<tr>
<th>Gene Target</th>
<th>Parasites</th>
<th>Primer Pairs</th>
<th>Reference</th>
<th>Primer Sequence</th>
<th>Ta (°C)</th>
<th>Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mt COI</td>
<td>Cystoisospora suis</td>
<td>COL_10F</td>
<td>Ogedengbe et al., 2011</td>
<td>GGWDSWGWRYWGGWTGGAC</td>
<td>50</td>
<td>517</td>
</tr>
<tr>
<td></td>
<td>Cystoisospora felis</td>
<td>COL_500R</td>
<td>Ogedengbe et al., 2011</td>
<td>CATRTGRGDCGCCAWAC</td>
<td>55</td>
<td>795</td>
</tr>
<tr>
<td></td>
<td>Hammondia heydorni</td>
<td>COL_10F</td>
<td>Ogedengbe et al., 2011</td>
<td>GGWDSWGWRYWGGWTGGAC</td>
<td>55</td>
<td>841</td>
</tr>
<tr>
<td></td>
<td>Hammondia hammondi</td>
<td>COL_780R</td>
<td>Ogedengbe et al., 2011</td>
<td>CCCAGAGATAATACAAAATGG</td>
<td>50</td>
<td>875</td>
</tr>
<tr>
<td></td>
<td>Toxoplasma gondii</td>
<td>COL_10F</td>
<td>Ogedengbe et al., 2011</td>
<td>GGWDSWGWRYWGGWTGGAC</td>
<td>55</td>
<td>841</td>
</tr>
<tr>
<td></td>
<td>Caryospora bigenetica</td>
<td>COL_1202R</td>
<td>El-Sherry et al., 2013</td>
<td>CCAAKRAYHGCACCAAGAGATA</td>
<td>50</td>
<td>875</td>
</tr>
<tr>
<td></td>
<td>Hammondia heydorni</td>
<td>Sdae-COI_260F</td>
<td>Present study</td>
<td>GATCTTTATGTTYTTRATGCC</td>
<td>50</td>
<td>875</td>
</tr>
<tr>
<td></td>
<td>Hammondia hammondi</td>
<td>Sdae-COI_1147R</td>
<td>Present study</td>
<td>CATTACCCATAACYACACC</td>
<td>50</td>
<td>875</td>
</tr>
<tr>
<td></td>
<td>Toxoplasma gondii</td>
<td>MEDLIN A</td>
<td>Medlin et al., 1988</td>
<td>AACCTGGTTGATCCTGCCAGT</td>
<td>50</td>
<td>875</td>
</tr>
<tr>
<td></td>
<td>Cystoisospora felis</td>
<td>MEDLIN B</td>
<td>Medlin et al., 1988</td>
<td>GATCTTTCTGACGTTACCTAC</td>
<td>50</td>
<td>875</td>
</tr>
<tr>
<td></td>
<td>Cystoisospora canis</td>
<td>Cystoisospora rivolta</td>
<td>Present study</td>
<td>CGAATAATCCACCGGAACACTCA</td>
<td>50</td>
<td>875</td>
</tr>
<tr>
<td></td>
<td>Cystoisospora rivolta</td>
<td>Sarco_18S_1711R</td>
<td>Present study</td>
<td>TATCAGCTTTGACGCTAGTGATT</td>
<td>50</td>
<td>875</td>
</tr>
<tr>
<td></td>
<td>Sarco_18S_123F</td>
<td>ERIB1OR</td>
<td>Barta et al., 1997</td>
<td>ACCTGGTTGATCCTGCCAG</td>
<td>50</td>
<td>875</td>
</tr>
</tbody>
</table>

11.3.5. PCR - PARTIAL mtCOI AND NEAR–COMPLETE NUCLEAR 18S rDNA

PCR amplification was carried out in an MJ Mini thermal cycler (Bio Rad, CA, USA) with reactions consisting of 1× PCR buffer (Invitrogen, Carlsbad CA, USA) supplemented with 4mM MgCl₂, 200 μM dNTPs, 0.5 μM of each of the amplification primers (Table 11.1), 2.5 U of Platinum® Taq Polymerase (Invitrogen) and 50–100 ng DNA template. Cycling conditions were initial denaturation at 96 °C for 5min followed by 35 cycles of 94°C for 30s, annealing at 50°C-55°C (see Table 1 for specific primer combinations) for 30s and extension at 72°C for 30-90 sec (depending on expected
product size). The PCR reaction was completed with a final extension step at 72 °C for 7–10 min. PCR reaction products were electrophoresed using a 1.5% submarine agarose gel, stained with ethidium bromide and visualized using UV trans-illumination (Spectronics Corporation, New York, USA). The apparent size of DNA bands were determined by comparison with a 100bp DNA ladder (Bio Basic Inc, Markham Ontario Canada). Bands were excised using a fresh scalpel blade and gel purified using a QIAGEN Gel Extraction Kit (Valencia, CA, USA) according to the manufacturer’s instructions. The purified PCR products were cycle sequenced using an ABI Prism 7000 Sequence Detection System (Applied Biosystems Inc., Foster City CA, USA) by the Molecular Biology Unit of the Laboratory Services Division, University of Guelph (Guelph ON, Canada) using the appropriate amplification primers in both directions. For the longer nuclear 18S rDNA fragments, internal sequencing primers were used so that complete, double-stranded sequencing was obtained.

11.3.6. SEQUENCE ASSEMBLY

The sequencing reads were assembled using de novo assembly within the Geneious bioinformatics software package. For both loci, the strict consensus sequence, less the amplification primers, was submitted to GenBank to obtain accession numbers reported below.

Added to newly generated sequences, existing publically available COI and corresponding complete or near–complete 18S rDNA sequences generated from the same strain or isolate, when available, were used preferentially for phylogenetic analyses (see Table 11.3 for sequences and strains used in this study). GenBank sequence FJ357797, previously identified as an Isospora sp. (strain Harbin/01/08) retrieved from the Siberian tiger (Panthera tigris altaica), was renamed as Cystoisospora sp. (strain Harbin/01/08) in this study on the basis of sporulated oocysts possessing sporocysts with Stieda bodies and with sporozoites lacking refractile bodies (see Figures 2A and 2B of Zhijun et al., 2011); these morphological features are consistent with the genus Cystoisospora rather than Isospora (see Barta et al., 2005). In addition, the following selected mt COI and 18S rDNA sequences from apicomplexan parasites outside of the Eimeriorina were used as the taxonomic outgroup taxa for rooting the resulting trees:
Hepatocystis sp.; Plasmodium juxtanucleare; P. vivax; P. falciparum; P. malariae; Babesia caballi; B. bovis; B. bigemina; B. rodhaini; Theileria annulata and T. parva. Accession numbers for all sequences included in the analyses are indicated on the resultant trees and Table 11.3.

Where possible, sequences obtained following bacterial cloning were excluded in the analyses to avoid sequence variations resulting from nucleotide misincorporation during PCR amplification (Olivieri et al., 2010; Ogedengbe et al., 2013). The 18S rDNA sequence from Isospora rivolta (AY618554) was removed from all analyses because a BLAST search determined that this sequence was identified incorrectly; AY618554 is likely derived from a basidiomycete fungus (see Whipps et al., 2012).

11.3.7. MULTIPLE SEQUENCE ALIGNMENT

Multiple sequence alignments based on the primary structure were generated for the partial mt COI sequences using the ‘Translation Align’ algorithm within Geneious Ver. 6.1.8 bioinformatics software (Cost Matrix=Blosum62; Gap open penalty=12; Gap extension penalty=3) (http://www.geneious.com, Kearse et al., 2012).

Nuclear 18S rDNA sequences were trimmed at both ends to exclude primer regions with sequences starting at the identical homologous nucleotide position 14 and ending at position 1778 if compared to the Genbank sequence DQ060683 of C. belli (Gjerde, 2013b). Alignments were performed first with MAFFT (Katoh et al., 2009; Katoh and Toh, 2010) and ClustalW (Larkin et al., 2007) executed from within Geneious. The occurrence of hypervariable regions in multiple sequence alignment of the 18S rRNA gene warranted that the MAFFT alignment was further staggered to address regions for which positional homology was uncertain (Ogedengbe et al., 2011). Non homologous residues were staggered, not aligned to other residues; distinct alignments blocks were then constructed through these regions thereby reducing the chances of introducing nonhomologous characters into the aligned data matrix (see Barta, 1997). The sequence alignments were analysed as three separate sequence datasets: 1) an ‘18S rDNA dataset’; 2) a ‘COI dataset’; and, 3) a ‘concatenated dataset’. The 18S rDNA dataset consisted of 100 aligned nucleotide sequences obtained from 61 spp. There were a total of 3273 character positions in the final 18S rDNA dataset of
which 1689 characters were constant, 904 informative characters and 685 variable characters were parsimony uninformative. The COI dataset consisted of 91 aligned nucleotide sequences obtained from 59 spp. There were a total of 756 character positions in the final dataset of which 226 characters were constant, 501 informative characters and 29 were parsimony - uninformative. The concatenated dataset combined the nuclear 18S rDNA and mt COI datasets into a single dataset without modifying the pre-existing sequence alignments for each genetic locus. The concatenated dataset had 4061 total characters of which 1943 characters were constant, 1408 parsimony – informative characters and 710 variable characters were parsimony uninformative. All positions containing gaps were treated as missing data.

11.3.8. PHYLOGENETIC ANALYSES

Phylogenetic analyses for all datasets (i.e. 18S rDNA, COI or concatenated) were performed using three tree building methods, Bayesian Inference (BI) executed from within Mr Bayes Version 3.1.2. (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003), Maximum Likelihood (ML) (Guindon et al, 2010) executed from within PAUP version 4.0 Beta (Swofford, 2002) and Maximum Parsimony (MP) (Felsenstein, 1985). Best fit models and parameters for the analyses were based on Akaike Information Criterion (AIC) for 24 models of DNA substitution using the hierarchical likelihood ratio test performed within MrModeltest v2.3 (Nylander J. A. A. 2004. MrModeltest v2 Program; distributed by the author, Evolutionary Biology Centre, Uppsala University). Based on AIC, the Hasegawa, Kishino and Yano (HKY) - Hasegawa et al., 1995) substitution model with a discrete gamma distribution including invariable site (G+I, 2 gamma categories) estimated base frequencies was recommended for the 18S rDNA data set. The general time reversible model with discrete gamma distribution including invariable site (GTR +I+ G, 6 gamma categories) was recommended for both the COI and the concatenated sequence data sets.

Bayesian Inference was performed with 2,000,000 generations of Markov Chain Monte Carlo with a sampling frequency of 1000 for all datasets. In the case of the ML analyses, bootstrap consensus tree was constructed from 500 replicates in all analyses. Branches with less than 50% bootstrap support were collapsed. Maximum likelihood
method was performed on the COI dataset with a gamma distribution shape parameter of 1.2495 and a rate category of 4; proportion of invariable sites (I), 0.2621; and number of distinct patterns under this model was 561. In COI MP analysis tree length was 3641 and consistency index (CI) was 0.3087. For the nuclear 18S rDNA dataset, shape parameter 0.2769 was used and a rate category of 4 was implemented; the proportion of invariable sites (I) was 0.1791 and the number of distinct data patterns under this model was 1580. MP values for the 18S rDNA analysis was 3100 for the tree length and consistency index (CI) 0.7200. Similarly, the number of distinct data pattern for the concatenated sequence data set was 2140; proportion of invariable sites (I), 0.2041; number of rate categories, 4 and gamma distribution shape parameter, 0.4366. Tree length in MP analysis was 6766 with consistency index (CI) 0.4957.

In all analyses, trees were rooted using Haemosporida and Piroplasmida (Aconoidasida). Only BI trees (Figures 11.2-11.4) are shown. Trees were drawn such that horizontal branch lengths are proportional to hypothesized genetic divergence.

11.4. RESULTS

11.4.1. OOCYST DIMENSIONS

Representative oocysts collected from pigs, cats or dogs are illustrated in Figure 11.1. The mean dimensions for oocysts and sporocysts of parasites isolated from fecal material are summarized for these parasites in Table 11.2.

11.4.2. SEQUENCES OBTAINED:

The primer pairs designed to amplify a portion of the mt COI gene (see Table 11.1) produced PCR products that range from 491bp to 888bp in length depending on the combination of amplification primers and excluding the primer regions. Final lengths for mt COI ranged from 418bp to 848bp. New sequences from nuclear 18S rDNA ranged from 573bp (partial sequence from C. rivolta) to near full-length sequences of ~1800bp. A total of 45 new COI sequences and 32 new 18S rDNA sequences were generated from 26 different species (see Table 11.3 for the GenBank accession numbers for all sequences used in this study; newly generated COI and 18S rDNA sequences and accession numbers are bolded).
Table 11.2. Oocyst and sporocyst dimensions for parasites obtained from fecal samples.

<table>
<thead>
<tr>
<th>Parasite Species (Isolate or Strain)</th>
<th>Host</th>
<th>n</th>
<th>Oocysts Length × Width (range) (µm)</th>
<th>Shape Index (range)</th>
<th>Sporocysts Length × Width (range) (µm)</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cystoisospora felis (MEO2010a-d07.2010.05.25)</td>
<td>Cat</td>
<td>25</td>
<td>41.1 (34-45) × 30.9 (28-34)</td>
<td>1.3 (1.5-1.1)</td>
<td>21.5 (18-25) × 18.3 (17-23)</td>
<td>25</td>
</tr>
<tr>
<td>Cystoisospora canis (MEO2015 IDX032)</td>
<td>Dog</td>
<td>19</td>
<td>35.1 (27-42) × 29.5 (22-32)</td>
<td>1.2 (1.4-1.0)</td>
<td>19.2 (15-22) × 14.4 (14-16)</td>
<td>8</td>
</tr>
<tr>
<td>Cystoisospora cf. ohioensis (MEO2014 Sedona)</td>
<td>Dog</td>
<td>25</td>
<td>24.2 (21-27) × 21.2 (17-24)</td>
<td>1.1 (1.3-1.0)</td>
<td>15.9 (13-26) × 12.3 (10-24)</td>
<td>25</td>
</tr>
<tr>
<td>Cystoisospora suis (MEO2015-Friendship7)</td>
<td>Pig</td>
<td>25</td>
<td>20.5 (13-23) × 17.6 (10-20)</td>
<td>1.2 (1.3-1.1)</td>
<td>11.4 (10-14) × 9.6 (8-11)</td>
<td>25</td>
</tr>
<tr>
<td>Cystoisospora rivolta (MEO2015 IDX 044)</td>
<td>Cat</td>
<td>25</td>
<td>22.4 (20-24) × 20.8 (18-24)</td>
<td>1.1 (1.3-1.0)</td>
<td>n/a^1</td>
<td>0</td>
</tr>
<tr>
<td>Hammondia heydorni (MEO2010e -AHL Braun)</td>
<td>Dog</td>
<td>25</td>
<td>12.2 (11-13) × 10.5 (9-12)</td>
<td>1.4 (1.5-1.2)</td>
<td>8.3 (8-9) × 6.1 (5-7)</td>
<td>25</td>
</tr>
</tbody>
</table>

^1n/a - sporulated oocysts were not available for this species.

Figure 11.1. Sporulated oocysts of some tissue coccidia. Two Cystoisospora species that infect dogs (A, Cystoisospora canis; B, Cystoisospora ohioensis), Two Cystoisospora species that infect cats (D, Cystoisospora felis; F, unsporulated Cystoisospora rivolta). Cystoisospora suis, infecting pigs (E) and Hammondia heydorni, a parasite of dogs (C).
<table>
<thead>
<tr>
<th>PARASITE</th>
<th>18S rDNA SEQUENCE ACCESSION NUMBERS</th>
<th>COI SEQUENCE ACCESSION NUMBERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hepatocystis sp.</td>
<td>EU400394 (Strain MFRC11)</td>
<td>EU254570 (Strain MB6)</td>
</tr>
<tr>
<td>Plasmodium falciparum</td>
<td>M19172</td>
<td>M76611</td>
</tr>
<tr>
<td>Plasmodium juxtamineforme</td>
<td>AF463507</td>
<td>AB250415</td>
</tr>
<tr>
<td>Plasmodium malariae</td>
<td>AB489196 (Strain Oumu)</td>
<td>AB489193 (Strain Oumu)</td>
</tr>
<tr>
<td>Plasmodium vivax</td>
<td>JQ627158 (Strain SV6)</td>
<td>AY791687 (PV02107b)</td>
</tr>
<tr>
<td>Babesia bigemina</td>
<td>AY603402 (Isolate Kuning strain Yunnan)</td>
<td>JQ518300 (Isolate Kuning)</td>
</tr>
<tr>
<td>Babesia bovis</td>
<td>HQ264112 (Isolate USDA A1)</td>
<td>NC_009902 (Strain 12Bo)</td>
</tr>
<tr>
<td>Babesia caballi</td>
<td>AY309955 (Spain-1)</td>
<td>AY499086</td>
</tr>
<tr>
<td>Babesia rodhaini</td>
<td>AB049999</td>
<td>AB624357 (Strain Australian)</td>
</tr>
<tr>
<td>Theileria annulata</td>
<td>KF559356 (Shanmenxia)</td>
<td>JQ518291 (Shanmenxia)</td>
</tr>
<tr>
<td>Theileria parva</td>
<td>HQ895972 (Strain B15e)</td>
<td>AB499089 (Isolate Muguga)</td>
</tr>
<tr>
<td>Sarcocystis alsodes</td>
<td>KF831274 (Isolate AaC2.2)</td>
<td>KF831245 (Isolate AaC2.2)</td>
</tr>
<tr>
<td>Sarcocystis alsodes trans</td>
<td>KF831276 (Isolate AaC1.21)</td>
<td>KF831251 (Isolate AaC1.21)</td>
</tr>
<tr>
<td>Sarcocystis cruzi</td>
<td>KC209740 (Strain B4.21)</td>
<td>KC209600 (Strain B4.21)</td>
</tr>
<tr>
<td>Sarcocystis gruenieri</td>
<td>EF056010 (Isolate Sg IRIN)</td>
<td>KC209617 (Isolate R31.4)</td>
</tr>
<tr>
<td>Sarcocystis gigantea</td>
<td>KC209733 (Isolate S1.1)</td>
<td>KC209601 (Isolate S1.1)</td>
</tr>
<tr>
<td>Sarcocystis hortii</td>
<td>KF8313294 (Isolate Aa36.6)</td>
<td>KF831323 (Isolate Aa36.6)</td>
</tr>
<tr>
<td>Sarcocystis neuna</td>
<td>KF854255 (Strain BR2012-TWC2013)</td>
<td>KF854272 (Strain BR2012-TWC2013)</td>
</tr>
<tr>
<td>Sarcocystis rangi</td>
<td>EF056011 (Isolate Sr IRIN)</td>
<td>KC209668 (Isolate R31.22)</td>
</tr>
<tr>
<td>Sarcocystis rangiferi</td>
<td>EF056016 (Isolate Sfr2IRin)</td>
<td>KC209674 (Strain R31.22)</td>
</tr>
<tr>
<td>Sarcocystis rileyi</td>
<td>HM185742 (Strain Europe)</td>
<td>KT184389 (Strain ME02015-08131)</td>
</tr>
<tr>
<td>Sarcocystis silesia</td>
<td>GU120092 (Rosenthal Duck 08131)</td>
<td>KT184389 (Strain ME02015-08131)</td>
</tr>
<tr>
<td>Sarcocystis sinensis</td>
<td>JN226122 (Isolate Cc11.4)</td>
<td>KC209687 (Isolate Cc11.4)</td>
</tr>
<tr>
<td>Sarcocystis taenata</td>
<td>KC209744 (Strain B4.9)</td>
<td>KC209694 (Strain B4.9)</td>
</tr>
<tr>
<td>Sarcocystis tarandii</td>
<td>KF8313277 (Isolate AaC1.7)</td>
<td>KF831236 (Isolate AaC1.7)</td>
</tr>
<tr>
<td>Sarcocystis tarandivulpes</td>
<td>GQ250969 (Isolate StrIRin)</td>
<td>KC209701 (Isolate R31.26)</td>
</tr>
<tr>
<td>Sarcocystis tenella</td>
<td>EF467656 (Isolate St1IRl)</td>
<td>KC209715 (Isolate R31.25)</td>
</tr>
<tr>
<td>Cystoisospora canis</td>
<td>KT184368 (Strain ME02015a-Milly C5)</td>
<td>KT184385 (Strain ME02015-Milly)</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>KT184363 (Strain ME02015b-Milly C6)</td>
<td>Same as above</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>KT184361 (Strain ME02015c-Milly C7)</td>
<td>Same as above</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>KT184360 (Strain ME02015d-Milly C11)</td>
<td>Same as above</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>KT184362 (Strain ME02015c-Milly C17)</td>
<td>Same as above</td>
</tr>
<tr>
<td>Cystoisospora sp. (small canine Cysto. cl. ohioensis)</td>
<td>AB519675 (Strain 2-MM)</td>
<td>None</td>
</tr>
<tr>
<td>Cystoisospora sp.</td>
<td>L76471</td>
<td>None</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>None</td>
<td>KF854256 (Strain ME02010a-25.05.10-007)</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>None</td>
<td>KF854257 (Strain ME02010b-09038458)</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>None</td>
<td>KF854258 (Strain ME02010c)</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>None</td>
<td>KF854259 (Strain ME02010d_IDX002)</td>
</tr>
<tr>
<td>Cystoisospora ohioensis</td>
<td>KT184364 (Strain ME02015-#7)</td>
<td>KT184386 (Strain ME02015-#7)</td>
</tr>
<tr>
<td>Cystoisospora ohioensis</td>
<td>KT184365 (Strain ME02010b_IDX005)</td>
<td>KF854261 (Strain ME02010b_IDX005)</td>
</tr>
<tr>
<td>Cystoisospora ohioensis</td>
<td>KT184359 (Strain ME02015a-Sedona)</td>
<td>KT184384 (Strain ME02015a-Sedona)</td>
</tr>
<tr>
<td>Cystoisospora ohioensis</td>
<td>None</td>
<td>KF854260 (Strain ME02010a_IDX008)</td>
</tr>
<tr>
<td>Cystoisospora cf rivolta</td>
<td>KT184366 (Strain ME02015c-#50)</td>
<td>None</td>
</tr>
<tr>
<td>Cystoisospora cf rivolta</td>
<td>KT184367 (Strain ME02015d-IDX443)</td>
<td>KT184383 (Strain ME02015-IDX443)</td>
</tr>
<tr>
<td>Isospora sp Harbin/01/08</td>
<td>FJ357797</td>
<td>None</td>
</tr>
<tr>
<td>Cystoisospora suis</td>
<td>KF854252 (Strain ME02010a_0224265)</td>
<td>KF854262 (Strain ME02010a_0224265)</td>
</tr>
<tr>
<td>Cystoisospora suis</td>
<td>KF854251 (Strain ME02010b_Sample B)</td>
<td>KF854263 (Strain ME02010b_Sample B)</td>
</tr>
<tr>
<td>Cystoisospora suis</td>
<td>KF854254 (Strain ME02013a_FA)</td>
<td>None</td>
</tr>
<tr>
<td>Cystoisospora suis</td>
<td>KF854265 (Strain ME02013b_FB)</td>
<td>None</td>
</tr>
<tr>
<td>Hyalomelkridia lieberkuehni</td>
<td>AF298623</td>
<td>None</td>
</tr>
<tr>
<td>Besnoitia besnoiti</td>
<td>AY833646</td>
<td>None</td>
</tr>
<tr>
<td>Nephroisospora eptesici</td>
<td>EU334134</td>
<td>None</td>
</tr>
<tr>
<td>Hammondia cf. triflittae</td>
<td>None</td>
<td>(NEW Pending GenBank submission)</td>
</tr>
<tr>
<td>Hammondia hammondii</td>
<td>KF854253 (Strain H.H-20)</td>
<td>KF854267 (Strain H.H-20)</td>
</tr>
<tr>
<td>Hammondia hammondii</td>
<td>KT184369 (Strain H.H-20)</td>
<td>KT184387 (Strain H.H-20 ME02015)</td>
</tr>
<tr>
<td>Toxoplasma gondii</td>
<td>AF086498 (H.H-34)</td>
<td>None</td>
</tr>
<tr>
<td>Toxoplasma gondii</td>
<td>None</td>
<td>KF854273 (Strain CTG)</td>
</tr>
<tr>
<td>PARASITE</td>
<td>18S rDNA SEQUENCE ACCESSION NUMBERS</td>
<td>COI SEQUENCE ACCESSION NUMBERS</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>-------------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>None</td>
<td>HM771689 (Strain GT1)</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>KF854274 (Strain GT1)</td>
<td>KF854275 (Strain MAS)</td>
</tr>
<tr>
<td>None</td>
<td>KF854276 (Strain PTG)</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>KF854277 (Strain TgCatBr5)</td>
<td>KF854278 (Strain TgCatBr64)</td>
</tr>
<tr>
<td>None</td>
<td>KF854279 (Strain TgCatGcCal)</td>
<td></td>
</tr>
<tr>
<td>L37415 (ME49)</td>
<td>HM771690 (Strain ME49)</td>
<td></td>
</tr>
<tr>
<td>EF472967 (RH)</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>L23381 (RH)</td>
<td>M97703 (RH)</td>
<td></td>
</tr>
<tr>
<td>Hammondia heydorni</td>
<td>JX220987 (clone 5 - BARON)</td>
<td>KF854271 (Strain ME20010c_Baron)</td>
</tr>
<tr>
<td></td>
<td>JX220986 (clone 3 - BARON)</td>
<td>KF854271 (Strain ME20010c_Baron)</td>
</tr>
<tr>
<td></td>
<td>GG994224 (Isolate HtCanis-2000-1-N)</td>
<td>JX73250</td>
</tr>
<tr>
<td>Neospora caninum</td>
<td>HM771688 (Strain NC-1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AJ271354 (Strain NC-PV1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KT184370 (Strain NC3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KT184388 (Strain NC3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KT184386 (Strain NC3)</td>
<td></td>
</tr>
<tr>
<td>Eimeria acervulina</td>
<td>U67115 (Attenuated E. acervulina H&amp;C)</td>
<td>HM771673 (Strain NC3)</td>
</tr>
<tr>
<td>Eimeria sp (pseudoparasite)</td>
<td>None</td>
<td>KF854266 (ME2010)</td>
</tr>
<tr>
<td>Eimeria brunetti</td>
<td>U67116.1 (Attenuated E. brunetti VA)</td>
<td>HM771675 (strain Guelph 80)</td>
</tr>
<tr>
<td></td>
<td>KT184337 (strain Guelph 80)</td>
<td>HM771675 (strain Guelph 80)</td>
</tr>
<tr>
<td>Eimeria maxima</td>
<td>U67117 (strain Guelph 84)</td>
<td>HM771684 (strain M6)</td>
</tr>
<tr>
<td>Eimeria praecox</td>
<td>FJ236370 (Shanghai)</td>
<td>JQ659301 (strain Guelph 10)</td>
</tr>
<tr>
<td></td>
<td>KT184351 (Strain MEO2015_Type A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KT184352 (Strain MEO2015_Type B)</td>
<td></td>
</tr>
<tr>
<td>Eimeria mitis</td>
<td>FR775307 (E. mitis, mivatis Variant 3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FR775302 (E. mitis, mivatis Variant 1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>U67118 (Strain FS 50)</td>
<td></td>
</tr>
<tr>
<td>Eimeria necatrix</td>
<td>KT184349 (strain Guelph 84)</td>
<td>HM771680 (strain Guelph 84)</td>
</tr>
<tr>
<td>Eimeria tenella</td>
<td>EF210325 (Shanghai)</td>
<td>HM771676 (strain USDA 80)</td>
</tr>
<tr>
<td></td>
<td>U67121 (Attenuated E. tenella Penn State)</td>
<td>HM771676 (strain USDA 80)</td>
</tr>
<tr>
<td></td>
<td>KT184354 (Strain USDA 80)</td>
<td>HM771676 (strain USDA 80)</td>
</tr>
<tr>
<td>Eimeria adenoeides</td>
<td>KC305180 (Line 5)</td>
<td>KC305176 (Line 4)</td>
</tr>
<tr>
<td></td>
<td>KC305176 (Line 4)</td>
<td>KC305192 (Line 4)</td>
</tr>
<tr>
<td>Eimeria meleagrimitis</td>
<td>FR745914 (strain KCH)</td>
<td>FR846202 (strain KCH )</td>
</tr>
<tr>
<td></td>
<td>KC346359 (Line 5)</td>
<td></td>
</tr>
<tr>
<td>Eimeria falciformis</td>
<td>AF080614 (REVISE HM771682.1(Strain Chob2))</td>
<td>REVISE HM771682.1(Strain Chob2)</td>
</tr>
<tr>
<td></td>
<td>KT184339 (strain Chob2A)</td>
<td>REVISE HM771682.1(Strain Chob2)</td>
</tr>
<tr>
<td></td>
<td>KT184340 (strain Chob2B)</td>
<td>REVISE HM771682.1(Strain Chob2)</td>
</tr>
<tr>
<td>Eimeria zuernii</td>
<td>AB769656 (strain Guelph 2007)</td>
<td>HM771687 (strain Guelph 2007)</td>
</tr>
<tr>
<td>Eimeria trichosuri</td>
<td>FJ236370 (Clone I)</td>
<td>JN192136 (strain Marco)</td>
</tr>
<tr>
<td>Cyclospora catenans</td>
<td>AF111183 (strain Guatemala and Haiti)</td>
<td>KF854210 (strain Cyclo CDC_2013)</td>
</tr>
<tr>
<td>Caryospora bigenetica</td>
<td>KT184333 (Strain MEO2015a)</td>
<td>KF858986 (strain MEO2015a)</td>
</tr>
<tr>
<td></td>
<td>KT184332 (Strain MEO2015b-MTZ07-12108)</td>
<td>KF858910 (strain MTZ_07-12108)</td>
</tr>
<tr>
<td></td>
<td>AF060975</td>
<td></td>
</tr>
<tr>
<td>Isospora sp</td>
<td>None</td>
<td>KC346356 (strain 1 JRB-2013)</td>
</tr>
<tr>
<td>Isospora gryphoni</td>
<td>KF854245 (UWO-2010)</td>
<td>KC346355 (UWO-2010)</td>
</tr>
<tr>
<td>Isospora greineri</td>
<td>KF648870 (Genotype A)</td>
<td>KF648868 (Genotype A)</td>
</tr>
<tr>
<td>Isospora superbusi</td>
<td>KF648871 (Genotype B)</td>
<td>KF648809 (Genotype B)</td>
</tr>
<tr>
<td>Isospora sp. ex. S. canaria</td>
<td>KT184337 (Strain MEO2015-xd24-09035458)</td>
<td>KF858103 (strain JRB2015-xd24-09035458)</td>
</tr>
</tbody>
</table>
11.4.3. PHYLOGENETIC ANALYSES BASED ON COI AND 18S rDNA SEQUENCES

All phylogenetic analyses using BI or ML methods with COI, 18S rDNA or concatenated sequences produced topologically similar trees that all supported the monophyly of the family Sarcocystidae (BI trees illustrated in figures 11.2-11.4). In all analyses, three clades were consistently supported within the family: 1) a clade containing all Sarcocystis spp. (Sarcocystinae) with pp ≥0.97 and bootstrap support of ≥99%; 2) a clade containing all Cystoisospora spp. (including Nephroisospora eptesici in the COI sequence-based tree only) with pp ≥0.99 and bootstrap support of ≥99%; and, 3) a clade containing all of the included parasites in the genera Hammondia, Neospora, and Toxoplasma.

Besnoitia besnoiti and Hyaloklossia lieberkuehni were determined to branch near the base of the Toxoplasmatinae in the 18S rDNA and concatenated dataset trees but there were no COI sequences available for these parasites. The position of N. eptesici varied depending on the genetic loci. In the tree generated from the COI dataset, the N. eptesici clustered within a clade with all Cystoisospora spp. In the tree generated from the 18S rDNA dataset, N. eptesici was found in the same clade with Toxoplasmal/ Hammondial/ Neospora along with B. besnoiti; H. lieberkuehni branched earliest amongst all taxa in the Toxoplasmatinae. However, in the concatenated sequence data set, N. eptesici formed a well-supported sister clade to the Cystoisospora spp.

In the analyses based on 18S rDNA and concatenated datasets (Figures 11.2 and 11.4, respectively) monophyly of both the subfamilies Toxoplasmatinae and Sarcocystidae was well supported. However, monophyly of the Toxoplasmatinae was not supported in trees generated from COI sequences alone (Figure 11.3) because the rooting position within the family Sarcocystidae was between Cystoisosporal/ Nephroisospora clade and the remaining members of the Toxoplasmatinae rather than between the Sarcocystis spp. and other members of the Sarcocystidae as was found in trees based on the other datasets (Figures 11.2 and 11.4).

The relationships among closely related species within the Toxoplasmatinae were less well resolved, particularly in trees based on the 18S rDNA (Figure 11.2) or concatenated datasets (Figure 11.4). For example, trees based on the 18S rDNA
Figure 11.2. Phylogeny of eimeriid coccidia (Eimeriidae and close relatives) and isosporoid coccidia (Sarcocystidae) based on the 18S rDNA dataset. Trees were generated using Bayesian inference and rooted using a number of haemosporinid parasites. The monophyly of the coccidia (eimeriid and isosporoid coccidia) was strongly supported. Within the coccidia, trees based on the 18S rDNA dataset supported monophyly of the eimeriid coccidia as well as the family Sarcocystidae and its two subfamilies, Toxoplasmatinae (PP=1.00) and Sarcocystinae (PP=0.97). However, relationships among species within the Toxoplasmatinae were not well resolved.
Figure 11.3. Enlarged portion of the 18S rDNA sequence based tree (see Figure 11.2 for complete tree) illustrating relationships among isosporid coccidia. *Hyaloklossia lieberkuehni* branched basally to well supported clades that contained all remaining members of the Toxoplasmataceae (all *Cystoisospora* spp. in one; the remainder in the second). Monophyly of individual *Cystoisospora* species was supported rarely; as an example, a single polytomy contained 4 *Cystoisospora* spp. from 4 different hosts. In the second clade, *Besnoitia besnoiti* and *Nephrilloispora eptesici* branched basally to a clade of poorly resolved taxa belonging to the genera *Hammondia*, *Toxoplasma* and *Neospora*. Although monophyly of each of *N. caninum* and *T. gondii* was supported (PP=0.99 and 0.71, respectively), sequences from individual *Hammondia* spp. did not form monophyletic clades and the genus *Hammondia* was paraphyletic as well.
Figure 11.4. Phylogeny of isosporoid coccidia based on the COI dataset using the same outgroups as in Figure 11.2; only the clade containing members of the family Sarcocystidae are illustrated with the branch leading to the other taxa indicated. Trees generated using Bayesian inference supported monophyly of the family Sarcocystidae. Three major clades were supported within the family Sarcocystidae: 1) a clade of Sarcocystis spp.; 2) Cystoisospora spp. plus Nephroisospora eptesici; and, 3) a clade of species in the genera Hammondia, Neospora or Toxoplasma. Monophyly of the subfamily Toxoplasmatinae was not supported in this tree. Hammondia species do not form a monophyletic group; Hammondia species using dogs as their definitive host grouped with N. caninum whereas Hammondia spp. using cats as their definitive host grouped with T. gondii.
Figure 11.5. Phylogeny of isosporoid coccidia based on the concatenated dataset (18S rDNA and COI sequences) using the same outgroups as in Figure 11.2: only the clade containing members of the family Sarcocystidae are illustrated with the branch leading to the other taxa indicated. Monophyly of the Sarcocystidae and each of its subfamilies, Sarcocystinae and Toxoplasmatinae was supported strongly (all PP=1.00). Besnoitia besnoiti and Hyaloklossia lieberkuehnii branched near the base of the Toxoplasmatinae group. Nephropoispora eipesici formed a well-supported sister clade to the Cystoisospora spp. Individual species within the clade containing Toxoplasma, Neospora and Hammondia spp. were well supported with the exception of the Hammondia spp. infecting canids. Additionally, monophyly of species in the genus Hammondia was not supported. As in the tree based on the COI dataset alone, Hammondia species using dogs as their definitive host grouped with N. caninum whereas Hammondia spp. using cats as their definitive host grouped with T. gondii.
dataset (Figure 11.2) did not support monophyly of any of the *Cystoisospora* spp. except *C. felis*. The remaining *Cystoisospora* spp. formed a poorly resolved, polychotomous multi-species cluster in which sequences from individual species did not form monophyletic clades.

In the trees based on COI sequences (Figure 11.4), the monophyly of many species within the Toxoplasmatinae was supported. However, *C. suis*, *C. cf. ohioensis* and *N. eptesici* sequences were found in an unresolved polyphagy basal to the remaining *Cystoisospora* species. In contrast, the remaining sequences from *Cystoisospora* spp. were hypothesized to be monophyletic based on COI sequence-based analyses. The *Cystoisospora* spp. that possess large, egg-shaped oocysts (i.e. *C. felis* and *C. canis*) formed a well-supported monophyletic clade (PP=0.95). Similarly, species of typically heteroxenous parasites infecting dogs and cats (species in the genera *Toxoplasma*, *Neospora* and *Hammondia*) were all found in a well-supported monophyletic clade in trees based on COI sequences (Figure 11.4). However, species in the genus *Hammondia* remained paraphyletic. The canid-infecting *Hammondia* spp. (i.e. *H. trifftae*, *H. heydorni*, and an unnamed *Hammondia* sp. from North America [i.e. strains ‘Baron’ and ‘Braun’]) were each monophyletic and formed a sister clade to *N. caninum* whereas the felid-infecting *H. hammondi* was sister taxon to *T. gondii*.

In the concatenated dataset analyses, sequences from canid-infecting *Hammondia* species were not resolved into monophyletic clades representing named species although they all formed a well-supported monophyletic clade (Figure 11.5). Again, as observed in the COI sequence-based analyses (Figure 11.4), parasites using canids as definitive hosts (i.e. *H. heydorni*, *H. trifftae* and *N. caninum*) and parasites using felids as definitive hosts (i.e. *H. hammondi* and *T. gondii*) each formed monophyletic clades; but the monophyly of *Hammondia* spp. was not supported.

11.5. DISCUSSION

Monophyly of all genera in the Toxoplasmatinae remains uncertain despite numerous phylogenetic analyses based on nuclear rDNA sequences and other loci (e.g. Mugridge et al., 2000; Morrison et al., 2004; Morrison, 2009). Complete mitochondrial genome sequencing of members of the phylum Apicomplexa revealed significant
diversity in genome organization and physical structure (e.g. Feagin, 2000; Gray et al., 2004; Hikosaka et al., 2010, Lin et al., 2011; Ogedengbe et al., 2013; 2014; 2015a; 2015b; Ogedengbe and Barta, 2015). Evaluating the diversity of apicomplexan mitochondrial genomes has provided the information necessary in designing applicable primers for the family Sarcocystidae.

In the present study, Toxoplasmatinae-specific PCR primers were used to obtain partial mt COI sequences with the goal of resolving evolutionary relationships among these tissue coccidia that could not be resolved using nuclear 18S rDNA sequences (e.g. Morrison, 2009). Culture-derived parasites of known identity as well as parasites obtained from naturally infected hosts were used to obtain mt COI and, in some cases, nuclear 18S rDNA sequences. The dimensions of coccidian oocysts (see Table 11.2) isolated from the feces of various vertebrate hosts were in agreement with previously published dimensions for these parasites (e.g. Nemeseri [1960] for C. canis, Dubey [1975a] for C. cf. ohioensis; Dubey [1979] for C. rivolta; Wenyon [1923] for C. felis, Biester and Murray [1934] for C. suis, Tadros and Laarman [1983] for H. heydorni). Unfortunately, no Frenkelia spp. was available for use in the present study.

Monophyly of nuclear 18S rDNA sequences from individual Cystoisospora species was not supported in the current study; only sequences from C. felis formed a monophyletic ‘species’ clade. Sequences from remaining Cystoisospora spp. formed an unresolved polytomy. Similarly, monophyly of Hammondia spp. was unsupported; 18S rDNA sequences from Hammondia spp. formed a polytomy with those of T. gondii and N. caninum. Morrison et al. (2004) concluded that nuclear 18S rRNA gene sequences provide sufficient phylogenetic signal for deep relationships but have insufficient signal to discern species-level relationships within the Eimeriidae and Sarcocystidae. In addition to insufficient phylogenetic signal, nuclear 18S rDNA sequences are problematic because they possess numerous indels (most likely within the single-stranded loops between the helices, Morrison et al, 2004) that complicate sequence alignments upon which phylogenetic hypothesis and tests of monophyly are built (Barta, 1997). Variability within some hypervariable regions (e.g. E21-1, E21-3 and E21-5) in the 18S rDNA sequences from taxa in the Sarcocystinae can also make alignments
problematic for these parasites (Morrison et al., 2004). Deleting the ‘problematic’ regions from phylogenetic analysis may compromise the phylogenetic support for the various taxonomic groups (Morrison et al., 2004). Barta, (1997) suggested a staggered alignment approach as a solution for retaining hypervariable regions while still maintaining positional homologies among those taxa for which positional homology can be reasonably assumed. Mugridge et al. (2000) similarly demonstrated the incongruities in resolving relationships based on phylogenetic analysis using the large subunit rRNA gene. This was evident in the unresolved placement or branching pattern in rDNA sequences for *C. felis* and *Sarcocystis* spp. within the subfamily Toxoplasmatinae irrespective of whether subsets of the large rDNA sequences or different phylogenetic methods were used.

In the current study, five *C. canis* clones from a single fecal isolate were sequenced; the cloned 18S rDNA sequences obtained failed to resolve into a monophyletic grouping in the phylogenetic analyses based on this locus. Based on pairwise sequence comparisons of alignments of these 5 sequences, the clones clustered into at least two distinct sets of paralogous 18S rDNA sequences. One pair of sequences (‘Type A’ – clones 11 and 17) had pairwise identity of 99.8% and clustered into a well-supported (PP=0.98) monophyletic clade. The second set of three sequences (‘Type B’ – clones 5, 6 and 7) had mean pairwise identity of 99.6% but did not resolve into a monophyletic clade in the 18S rDNA analyses. All *C. canis* clone sequences, together with other *Cystoisospora* spp., formed a well supported clade for the genus *Cystoisospora* in the 18S rDNA (PP=1.00) and concatenated dataset (PP=1.00) trees.

Paralogous 18S rDNA loci in the nuclei of coccidian parasites are a known issue that complicates phylogenetic inference. Two types of 18S rDNA sequences was reported by El-Sherry et al. (2013) in a well-supported *E. meleagrismitis* clade; two distinct clades (Type A with five sequences and Type B with 11 sequences) were generated using 18S rDNA sequence data. Vrba et al. (2011) also reported two types of nuclear 18S rDNA in a single-oocyst cloned line of *E. mitis*. More broadly in the Apicomplexa, McCutchan et al. (1988) described paralogous transcripts in several *Plasmodium* spp that originated from distinct sets of ribosomal genes expressed at
different stages of the parasite life cycle (Li et al., 1997). In addition to the risk of paralogous genes, the low level of divergence in 18S rDNA sequences reduces the resolving power of the nuclear 18S rDNA sequences to differentiate closely related species within the family Sarcocystidae as illustrated in the present study and previously (Ogedengbe et al., 2011). Therefore, complete 18S rDNA sequences may not be appropriate to be used alone as molecular markers for phylogenetic analyses for the family Sarcocystidae.

Although Cystoisospora spp. have been included in the subfamily Toxoplasmatinae, Cystoisospora and Sarcocystis spp. grouped together in the tree based on the COI dataset. One possibility for this unexpected result could be an artifactual rooting point resulting from Long Branch attraction (Townsend et al., 2012) among the distantly related major clades within the family Sarcocystidae. Including additional biologically diverse members of the subfamily Toxoplasmatinae not included in the current COI dataset, such as Besnoitia or Hyaloklossia spp., might help stabilise the rooting point for taxa within the Toxoplasmatinae. Addition of some ‘primitive’ eimeriid coccidia, such as Goussia or Choleoeimeria spp., to the taxonomic outgroup might also promote stability of the rooting point within the Sarcocystidae for this genetic locus.

Hammondia species are paraphyletic in all analyses containing COI sequences and paraphyly was not refuted by the 18S rDNA dataset (e.g. Morrison et al., 2004). All canid-infecting species (i.e. Hammondia heydorni and unnamed Hammondia spp and H. trifrittæ) formed a monophyletic clade that was the sister group to N. caninum (see e.g Ellis et al., 1999). Sequences obtained from a canine isolate of a Hammondia spp. from central Canada was nearly identical with those of H. trifrittæ at both the COI and 18S rDNA loci suggesting that the range of H. trifrittæ may include the entire temperate north. Although closely related, H. heydorni and H. trifrittæ exhibit genetic and biological differences (Gjerde and Dahlgren, 2011; Schares et al., 2002, 2003; Abel et al., 2006).

Hammondia hammondi clustered with T. gondii in all analyses. Partial mt COI sequences from strains of T. gondii (CTG, GTI, MAS, ME49, PTG, TgCatBr5, TgCatBr64, TgCgCal, and TgToucan) were all identical. The mt COI sequence from the
H.H-20 strain of *H. hammondi* shared 98.5% identity to the *T. gondii* sequences. The 1.5% divergence between these two species is identical to the divergence between *C. felis* and *C. rivolta* (1.5%) and similar to divergence between two *Isospora* spp. infecting the same avian host (1.3%, Hafeez et al., 2014) or between the chicken parasites *E. tenella* and *E. necatrix* (1.7%, Ogedengbe et al., 2011). The genetic divergence supports species-level separation of *H. hammondi* from *T. gondii* in contrast to Mehlhorn and Heydorn’s (2000) and Heydorn and Mehlhorn’s (2001) argument against the recognition of *H. hammondi* as a distinct species. These authors suggested that *T. gondii* and *H. hammondi* may be strains of a single species based on oocyst wall structure, tachyzoite ultrastructure and molecular similarities (see Mehlhorn and Heydorn, 2000) and cross-immunity (see Frenkel and Dubey, 2000).

Phylogenetic analyses based on LSU rDNA and ITS1 sequence data (Ellis et al., 1999), similarities in G+C content of DNA sequences (Johnson et al., 1987), and oocyst structure and antigenic responses (Dubey et al., 2002a) have consistently indicated that species in the genus *Hammondia* do not form a monophyletic group. In the present study, the COI- and the concatenated sequence-based trees show that relationships among species within the subfamily Toxoplasmatinae reflected the definitive hosts (DHs) infected by the parasites. For example, *H. hammondi* and *T. gondii* that both infect felid DHs formed a monophyletic clade; similarly, *Hammondia* spp. and *N. caninum* that all infect canid DHs formed a monophyletic clade. In the 18S rDNA based tree, *H. heydorni*, *H. triffittae* and *N. caninum* formed a polytomy with *T. gondii* and *H. hammondi*.

Mehlhorn and Heydorn, (2000) proposed that the generic assignments be radically simplified. They suggested that *Toxoplasma gondii* and *H. hammondi* should be synonymised under a single species, *T. gondii*. The two described *Neospora* spp., *N. caninum* and *N. hughesi*, were to be synonymised under a single species, *H. heydorni*, and then *H. heydorni* should be transferred to the genus *Toxoplasma* to give *T. heydorni* as the name for these parasites of dogs. Synonymizing *N. caninum*, *N. hughesi* and *H. heydorni* under *T. heydorni* by Mehlhorn and Heydorn (2000) would appear to be a taxonomic lumping too far. The question is whether this arrangement reflected biological reality and the relationships reflected by COI sequences. Synonymizing *T.
gondii and H. hammondi under T. gondii does reflect the results of COI phylogeny. The COI-based tree resolved T. gondii and H. hammondi as a monophyletic group. In the present study, there were 10 single nucleotide differences (SND) between the H.-H20 strain of H. hammondi and the ME49 strain of T. gondii over 751bp of the COI gene (1.3% genetic difference); to put this in perspective, C. felis and C. canis have only 0.9% sequence divergence (6 SND) in the same region of the COI gene and they do not share the same definitive host.

The suggestion of Mehlhorn and Heydorn (2000) that the generic assignment of H. hammondi be examined critically is worth considering. However, to preserve the monophyly of species in the genus Hammondia, H. hammondi could be transferred to the genus Toxoplasma and thus re-classified as Toxoplasma hammondi to reflect its close taxonomic and biological similarity to T. gondii. Such a taxonomic revision will require assigning a new name-bearing type (a new genotype) for the genus Hammondia; Hammondia heydorni would seem to be a suitable species for this purpose. The remaining Hammondia spp. (all parasites of canid definitive hosts) would then form a monophyletic sister group to N. caninum and, presumably, N. hughesi for which mt sequence data were not available. In addition to biological differences (e.g. Dubey et al., 2002b) and other molecular studies (e.g. ITS - Ellis et al., 1999, LSU rDNA - Mugridge et al., 1999, alpha tubulin - Siverajah et al., 2003), the distinctiveness of N. caninum from H. heydorni (and likely H. trifittae – Gjerde and Dahlgren, 2011) using mt COI sequences was confirmed in the present study.

11.6. CONCLUSIONS

Identification of Cystoisospora spp. is complicated by the wide range of oocyst measurements previously reported for described species (Dubey and Frenkel, 1972a; Frenkel and Dubey, 1972; Fayer and Frenkel, 1979; Agrawal et al., 1981; Bjork et al., 2000). In particular, all members of the Cystoisospora ohioensis spp. complex (C. ohioensis Dubey 1975a, C. neorivolta Dubey and Mahrt 1978 and C. burrowsi Trayser and Todd 1978, parasites of canid hosts) produce morphometrically similar oocysts that make assignment to a particular species based on fecal oocyst morphometrics tenuous at best. Although partial COI sequences appear well-suited for species delimitation
(Ogedengbe et al., 2011) and inferring phylogenetic relationships among the eimeriid coccidia (Ogedengbe et al., 2011 and current study), the COI sequences from Cystoisospora spp. demonstrated limited sequence variation. For example, the partial COI sequence of C. cf. ohioensis (Sedona) differed from the COI sequences of C. suis and C. rivolta by 1 single nucleotide difference (SND) while differing from a second C. cf ohioensis (IDX005) isolate by 3 SNDs. It is possible that the pairwise differences in the COI sequences of the 2 C. cf. ohioensis isolates is detection of species level differences within parasites in the C. ohioensis spp. complex. Only isolation of a pure line of each parasite and observations on the endogenous development of these parasites would be able to assign these COI sequences unequivocally to C. ohioensis, C. burrowsi or C. neorivolta.

Phylogenetic trees based on both the COI and concatenated datasets supports H. hammondi as a sister species to T gondii. Unfortunately, the support for the close relationship of the other Hammondia spp. to N. caninum is equally strong. This confirms the lack of monophyly of species currently placed in the genus Hammondia as reported previously (Ellis et al., 1999; Johnson et al., 1987; Dubey et al., 2002a). Transferring H. hammondi into the genus Toxoplasma and assigning a new genotype for the genus Hammondia (perhaps H. heydorni) would appear to be the simplest taxonomic change that reflects both biological and molecular features of these parasites.

Relationships among Neospora, Toxoplasma and Hammondia spp. appear to reflect the definitive host used by each parasite; parasites using canine definitive hosts clustered together as did parasites infecting felids as definitive hosts. Obtaining mt COI sequences from other tissue coccidia, such as Besnoitia spp., and more distantly related isosporoid coccidia such as Hyaloklossia lieberkuehni may help to clarify this situation.
12. THE PARAPHYLETIC EIMERIID GENERA EIMERIA AND ISOSPORA: MORPHOTYPES CONFLICT WITH MOLECULAR PHYLOGENIES INFERRRED USING NUCLEAR (18S rDNA) AND MITOCHONDRIAL (COI) SEQUENCES

12.1. ABSTRACT

The speciose genus *Eimeria* (about 1800 described species) is paraphyletic. Mitochondrial (mt) cytochrome *c* oxidase subunit I (COI) and nuclear (nu) 18S rDNA sequences were obtained from well characterized single-oocyst clonal lines of *Eimeria* species infecting turkeys and chickens. New and published sequences for *Eimeria* species and related eimeriid coccidia in the genera *Cyclospora*, *Isospora*, *Caryospora* and *Lankesterella* were used in phylogenetic reconstructions based on nu 18S rDNA, mt COI and combined sequence datasets. Most analyses showed that parasites infecting closely related definitive hosts formed monophyletic clades except *Eimeria* species infecting poultry or rodents. Phylogenetic analyses did not support the hypothesis that *Eimeria* spp. infecting turkeys or chickens originated from a common ancestor; indeed, *Cyclospora cayetanensis* was consistently included within a clade of *Eimeria* species infecting galliform birds. *Eimeria dispersa* and *E. innocua* formed a clade distinct from the remaining *Eimeria* species infecting galliform birds. *Eimeria necatrix* and *E. tenella* (that infect chickens) were related to *Eimeria* species infecting turkeys more closely than to the other *Eimeria* spp. of chickens. Neither ‘*Eimeria*-type’ nor ‘*Isospora*-type’ oocyst morphotypes formed monophyletic groups suggesting that this fundamental generic feature may be a less reliable taxonomic marker than previously assumed.

12.2. INTRODUCTION

Coccidiosis of galliform birds is common and cosmopolitan (Chapman et al., 2013). Seven species of *Eimeria* are responsible for inducing coccidial infections in turkeys (Chapman, 2008) and a further seven or more species can infect chickens (Reid and Long, 1979; Barta et al., 1997). *Eimeria* species belong to the family Eimeriidae Minchin 1903. This biologically diverse family includes a large number of parasites that belong to various described genera including *Caryospora*, *Cyclospora*, *Eimeria* and *Isospora* (see Chapman et al., 2013) *Epieimeria* (see Dyková and Lom, 1981). The speciose genus *Eimeria* contains coccidia with many morphological and molecular
variations. Molecular phylogenies consistently fail to group all sampled *Eimeria* species into a single monophyletic group, suggesting strongly that the genus as currently defined is paraphyletic (Jirků et al., 2002; Chapman et al., 2013).

Although *Eimeria* species of chickens have been studied thoroughly and can be identified readily biologically and using molecular methods, this is not the case for *Eimeria* species that infect turkeys. Throughout the 20\(^{th}\) century, *Eimeria adenoeides, E. dispersa, E. gallopavonis, E. meleagrimitis, E. meleagris, E. innocua* and *E. subrotunda* were described from turkeys (Chapman, 2008). Three of these, *E. adenoeides, E. gallopavonis* and *E. meleagrimitis*, are considered the most important *Eimeria* spp. infecting turkeys (Long et al., 1987). The economic cost of coccidiosis to commercial poultry is considerable, with millions of dollars expended annually for control or lost through reduced production (Clarkson and Gentles, 1958; Chapman et al., 2013). Unlike use of in-feed anticoccidial medications, the use of live vaccines for coccidiosis control may require species-level identification of *Eimeria* species infecting poultry. Although *Eimeria* species produce oocysts that have relatively consistent shapes and sizes, these easily accessible diagnostic stages have dimensions that overlap significantly among parasites even within a single host (e.g. El-Sherry et al., 2015). Detecting infection of turkeys by *Eimeria* species is easy but identification of the specific species implicated in the infection is considerably more difficult. Without molecular methods, identification can be laborious (Poplstein and Vrba, 2011; Chapman et al., 2013; El-Sherry et al., 2013; 2015). The nuclear 18S rDNA loci have been used widely for molecular characterization and phylogenetics of various coccidia (e.g. Barta et al., 1997; Carreno and Barta, 1999; Morrison et al., 2004; Morrison, 2009). However, nuclear 18S rDNA sequences are of limited use in inferring relationships among closely related coccidia because of insufficient sequence variation between species (Barta et al., 1997; Li et al., 1997; Heger et al., 2011), and the presence of multiple, divergent gene copies (paralogs) within the nuclear genomes of apicomplexan parasites (McCutchan et al., 1988, 1995; Kibe et al., 1994; Lew et al., 2003; Goethert et al., 2006). Frequently, paralogous rDNA copies within a single parasite have greater sequence divergence than homologous rDNA copies from closely related parasites. Paralogous rDNA copies with substantial sequence diversity had been reported in single-oocyst clonal lines of some
Eimeria species (i.e. Eimeria mitis and Eimeria meleagrimitis; see Vrba et al., 2011 and El-Sherry et al., 2013, respectively). Alignment of 18S rDNA sequences obtained from more distantly related coccidia is complicated further by extensive indels that make assignment of positional homology in sequence alignments difficult or impossible (e.g. Barta, 1997; Morrison et al, 2004; Ogedengbe et al., 2011). More rapidly diverging sequences in other portions within the nuclear ribosomal gene arrays such as the internal transcribed spacer (ITS) regions are similarly subject to intragenomic sequence variation (Alverson and Kolnick, 2005). The variable internal transcribed spacer regions [ITS-1 and ITS-2] demonstrate considerable genetic diversity among strains of the same species and even within a single parasite (Cook et al., 2010). Their use for inferring phylogenetic relationships among species is therefore limited (Lew et al., 2003).

According to Hebert et al. (2003a) and Zhao et al. (2012) the mitochondrial cytochrome c oxidase subunit I (mt COI) gene is a comparatively quickly evolving genetic locus with a sequence divergence rate that may be 1% to greater than 2% per million years (Myr) that is well suited for delimiting closely related metazoan taxa; although sequence divergence rate is thought to be higher for protistan organisms (Hebert et al., 2003b). Partial (600 to 800bp) mt COI sequences from Eimeria species possess considerable interspecific sequence diversity and comparatively little intraspecific sequence variation (Ogedengbe et al., 2011). For this reason, COI sequences are highly effective for identification and differentiation of closely related eimeriid coccidia such as Eimeria species compared to complete or near-complete nuclear 18S rDNA sequences (Ogedengbe et al., 2011). Consequently, COI sequences have become useful genetic markers for Eimeria species infecting chickens (Ogedengbe et al., 2011; Vrba et al., 2011), rodents and rabbits (Kvičerová and Hypša, 2013), and turkeys (El-Sherry et al., 2013; 2015).

In this study, we used sequence data from the mt COI and nu 18S rRNA genes to infer phylogenetic relationships of Eimeria species infecting turkeys and other galliforms among a range of eimeriid coccidia infecting a range of vertebrate hosts.
12.3. MATERIALS AND METHODS

12.3.1. OOCYSTS SOURCE, PURIFICATION AND DNA EXTRACTION

Oocysts of parent stocks for *E. adenoeides* Weybridge, *E. adenoeides* USKS06-01 strain, *E. meleagridis* USAR97-01 strain, *E. dispersa* Briston strain and *E. meleagrimitis* USMN08-01 were kindly provided by Dr. H.D. Chapman (Department of Poultry Science, University of Arkansas, Fayetteville AR, USA). *Eimeria adenoeides* Guelph strain was obtained in 1985 from a commercial turkey flock in Ontario, Canada. Single species lines from all parent stocks were derived from single oocysts *in vivo* through propagation in parasite free poults at the Central Animal Facility (CAF) Isolation unit, University of Guelph, Guelph ON, Canada. The method of Remmler and McGregor (1964), with slight modification, was used to generate single lines for each species by delivering individual agar plugs carrying a single oocyst within gelatin capsules to the crop of birds. For this study, progeny of derived single lines for all species were used. A detailed description of the origins of the original isolates from which each line was derived is provided by El-Sherry et al. (2015).

Oocysts of laboratory strains of various poultry *Eimeria* species were available for study, including: *E. brunetti* Guelph 80, *E. acervulina* NC3, *E. maxima* M6, *E. maxima* Guelph 74, *E. necatrix* Guelph 84, *E. praecox* Guelph 2012, *E. tenella* USDA 80. Additional oocysts were purified from diagnostic fecal specimens from the Animal Health Laboratory, Laboratory Services Division, University of Guelph (Guelph ON Canada) or from field samples. Parasites isolated from clinical or field fecal samples include: *Eimeria alabamensis*, *E. bovis* and *E. zuernii* from domestic cattle; *E. ahsata* from domestic sheep; *E. tamiasciuri* from the American red squirrel (*Tamiasciurus hudsonicus*); *Caryospora bigenetica* from a Massasauga rattlesnake (*Sistrurus catenatus*) (see Ogedengbe and Barta, 2015 for details); and, *Isospora gryphoni* from the American Goldfinch (*Carduelis tristis*). Infected tissues containing an *Isospora* sp. causing visceral coccidiosis in canaries (*Serinus canaria*) was provided by the Animal Health Laboratory (see Ogedengbe et al., 2015 for further details). Amphibian blood containing sporozoites of *Lankesterella minima* were obtained from *Lithobates clamitans*. We thank Dr. Yvonne Qvarnstrom (United States Centers for Disease
Control and Prevention, Atlanta, GA, USA) for supplying DNA templates for *Cyclospora cayetanensis* and Dr Bill Chobotar (Andrews University, Berrien Springs MI USA) for supplying oocysts of *Eimeria vermiformis* Chob 3, *Eimeria falciformis* Chob 2, and *Eimeria papillata* Chob 1. The laboratory strain (NC1) of *Neospora caninum* was a kind gift of Dr. J.P. Dubey (United States Department of Agriculture, Beltsville MD, USA).

Purification of oocysts from fecal material and subsequent genomic nucleic acid extraction was carried out as previously described by El-Sherry et al. (2013) and Ogedengbe et al. (2013). Disruption of oocysts was aided by the use of glass beads and vortexing using a Fisher vortex mixer (Fisher Scientific Inc, Bohemia N.Y 11716, USA). DNA extraction was performed using DNAzol (Invitrogen, Carlsbad, CA) as directed by the manufacturer. DNA was extracted from tissues or blood similarly but without the use of glass beads. DNA concentration was estimated by spectrophotometer (Nanodrop 2000, NanoDrop Products, Wilmington DE, USA). DNA was stored at 4°C for immediate use or -20°C for later use.

12.3.2. PCR of Nuclear 18S rDNA and Mitochondrial COI Genes

Near-complete, nuclear small subunit ribosomal DNA (18S rDNA) was amplified using two sets of universal eukaryote-specific primers: 1) Medlin A and Medlin B (Medlin et al., 1988; and, 2) ERIB_1F and ERIB_10R (Barta et al., 1997). Two alternate amplification primers 18S_224F (5’-TCATAGTAACCGAACGGATC-3’, paired with ERIB_10R) and CPB-DIAGR2 (5’-TAAGGTTGCTGAAGGAGTAAGG-3’, Johnson et al., 1995, paired with Medlin A) were used to obtain partial 18S rDNA sequences from *L. minima* and *E. vermiformis* respectively. Similarly, 475bp to 761 bp COI sequences were newly generated for all other coccidian parasites used in the current study using various combinations of degenerate COI primers (i.e. 10F (5’-GGWDWGGWRYWGGGTGGAC-3’) / 500R (5’-CATRTGRTGDCCCAWAC-3’) or 400F (5’-GGTCGGTGGTTGTGGAC-3’) / 1202R (5’-CAAKEYHGCACCAAGGATA-3’) or 10F / 1202R). The forward degenerate primers were concatenated with the universal M13Forward primer and the reverse primers with M13Reverse primer for ease of subsequent sequencing of PCR amplicons. PCR amplification was carried out in an MJ Mini thermal cycler (Bio-Rad, Hercules, CA,
USA) and reactions consisted of 2.5 U of Platinum® Taq Polymerase (Invitrogen, Burlington, ON, Canada), 1× PCR buffer (Invitrogen, Carlsbad CA, USA) supplemented with 4mM MgCl₂, 200μM of each dNTP, 0.5 μM of each of the amplification primers and 50-100 ng DNA template. The PCR thermal cycling profile consisted of initial denaturation at 96°C for 5min followed by 35 cycles of 94°C for 30s, annealing at 50-55°C (depending on primer combination) for 30s and extension at 72°C for (60-90s). The PCR was completed with a final extension step at 72 °C for 10min (for both 18S rDNA and COI). Negative and positive control templates were included in the reactions. PCR amplicons were electrophoresed at 80V for 45 min through a 1.5% submarine agarose gel containing ethidium bromide in 1×Tris-Acetate-EDTA (TAE) buffer. The resulting stained gel was visualized using UV transillumination (Spectronics Corporation, New York, USA). The apparent size of DNA bands were determined by comparing bands to a 100 bp - 10 kb DNA ladder standards (Bio Basic Inc., Mississauga ON, Canada). Excised gels containing bands were purified using a QIA quick gel extraction and purification kit (Qiagen, Toronto ON, Canada) according to the manufacturer’s instructions. The purified PCR products were cycle sequenced with both forward and reverse primers using an ABI Prism 7000 Sequence Detection System (Applied Biosystems Inc., Foster City CA, USA) by the Molecular Biology Unit of the Laboratory Services Division, University of Guelph (Guelph ON, Canada). In order to generate near-complete 18S rDNA sequences, four internal sequencing primers were used in addition to the PCR amplification primers: 18S Cocci_1205F (5’-GGTTCTGGGGGGAGTATGG-3’); 18S Cocci_1252R (5’-AAYTCCTTTAAGTTTCAG-3’); 18S_676F (5’-GTTGCAGTTARAARGCTCGTA-3’) and 18S_696R (5’-TACGAGCYTTYTAACTGCAAC-3’).

Chromatograms of newly generated sequences were assembled using the de novo sequence assembler embedded in the bioinformatics software package Geneious Version 6.1.8 (http://www.geneious.com, Kearse et al., 2012). Amplicons were sequenced completely in both directions to obtain high quality contigs for both 18S rDNA and mt COI PCR products. Primer regions were trimmed from newly generated sequences prior to GenBank submission.
12.3.3. DATA TREATMENT AND MULTIPLE SEQUENCE ASSEMBLY

Three datasets were analysed in the present study: 1) near complete 18S rDNA sequences; 2) partial mt COI sequences; and, 3) a concatenated dataset containing both nu 18S rDNA and mt COI sequences. Publically available mt COI and nuclear 18S rDNA sequences from eimeriid coccidia were retrieved from GenBank and combined with newly generated sequences within each dataset. For some parasites, large numbers of sequences have been submitted for each of these genetic loci. To reduce the absolute number of individual sequences to be analysed and as a means of generating a representative sequence for each genetic locus for each species, consensus sequences were generated from available sequences. For a single parasite species with ≥10 GenBank sequences (either newly submitted or previously deposited to GenBank) at a particular genetic locus, a 90% majority rule consensus sequence was generated for that locus and for that parasite. Similarly, for a single parasite species with <10 available GenBank sequences at a locus, a 50% strict consensus sequence was generated. The GenBank accession numbers of sequences from which the consensus sequences were derived for each genetic locus can be found in the supplementary material (Appendix 4 Supplementary Table 12.1).

The aligned 18S rDNA sequences from several parasites demonstrated divergent sequences that clustered into two apparent ‘types’ that were interpreted to represent paralogous 18S rDNA loci (sensu El-Sherry et al., 2013). The a priori assumption was made that these divergent clusters represented paralogs within the nuclear genome and were thus treated separately as Type A and Type B for each parasite for which this was observed. In these cases, the appropriate 50% or 90% consensus sequence for each type was generated as described above, depending on the number of each ‘type’ available.

Excluded in the current analyses were 18S rDNA sequences interpreted to be chimeric PCR products resulting from the amplification of DNA from mixtures of Eimeria species infecting chickens or turkeys as previously discussed (El-Sherry et al., 2013). Accession numbers for the excluded sequences are EU044768; EU044770; EU044771; EU044772; EU044774; EU044775; EU044776; EU044777; EU044778; EU044783; EU044784; HM117011; HM117015; HM117016 and EF210326. In
addition, KC305188 and KC305198 (*E. meleagrimitis* 18S rDNA sequences) that appeared to be PCR chimeras of Type A and Type B rDNA paralogs were excluded for the same reason. However, AF041437 (*E. meleagrimitis* 18S rDNA) was excluded from analyses because of low sequence quality.

Initial alignment for all sequences (18S rDNA and mt COI) was based on the primary structure using the Multiple Align algorithm within Geneious (Cost Matrix=65% similarity; Gap open penalty=12; Gap extension penalty=3). For the 18S rDNA sequence dataset, multiple alignment was performed with MAFFT (Katoh and Toh, 2010) executed from within Geneious. Alignments were inspected visually to correct obvious alignment errors manually (Templeton et al., 2010). Positions containing gaps were treated as unknown. In the COI sequence dataset, open reading frames were determined and confirmation of translation were performed using the mold/protozoan mitochondrial codon translation (i.e. translation table 4) within the Geneious bioinformatics package (Templeton et al., 2010) as described (Ogedengbe et al., 2014).

The third dataset consisted of concatenated 18S rDNA and mt COI consensus sequences. The concatenated dataset was created by concatenating the two previous datasets without modifying the alignments within each gene.

12.3.4. PHYLOGENETIC ANALYSES

Phylogenetic analyses of each of the genetic loci (i.e. the 18S rDNA or COI datasets) were based on two tree-building methods. Analyses were performed with the Bayesian Inference (BI) (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck, 2003), Maximum Likelihood (ML) (Guindon et al., 2010) methods.

The Best fit models and parameters for the analyses were selected based on Akaike Information Criterion (AIC) for 24 models of DNA substitution using the hierarchical likelihood ratio test, evaluated in MrModeltest v2.3 (Nylander, 2004). MrModeltest v2.3 Program; distributed by the author, Evolutionary Biology Centre, Uppsala University). Optimal Best fit model selected for BI and ML analyses for the nuclear 18S rDNA dataset was the General Time Reversible (GTR) model (Tavaré, 1986), with invariant gamma (I+G), distribution rate at 6 variation categories of nucleotide substitution. The same model was used for the COI dataset with the exception
that Codon nucleotide model (i.e. Nucmodel=Codon) and the translation table for the
codon analysis was changed to metazoan mitochondrial translation (i.e. Code=metmt).
Missing data was represented and treated as unknown. The concatenated dataset was
partitioned into two and the GTR model with invariant gamma (I+G), distribution rate at
6 variation categories of nucleotide substitution was executed for the 18S rDNA portion
of the combined dataset, while the Nucmodel with mitochondria translation
Code=metmt was implemented for the COI portion of the dataset.

The Bayesian Inference method was executed from MrBayes bioinformatic
software (Version 3.1.2.) (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck,
2003). Analyses under the general time reversal (GTR+I +G) consisted of 1×10^6
generations of Markov Chain Monte Carlo searches containing four chains, heated chain
temperature of 0.2 and burn-in length at 200,000. Subsample frequency of 1,000 for all
datasets ensured that trees were saved every 1,000 generation.

The ML analyses (Guindon and Gascuel, 2003), was performed using PhyML
(Guindon et al., 2010) executed from within the Geneious bioinformatics software
package (Version 6.1.8). The ML bootstrap analysis from 500 replicates was performed
to estimate node support. Branches with less than 50% bootstrap support were collapsed.
Gaps were treated as unknown or missing characters in all datasets. In the ML model
for the 18S rDNA datasets, shape parameters 0.4414 and 0.7125 and proportion of
invariable sites (I) of 0.6133 and 0.3650 were implemented, respectively.

The phylogenetic analyses assemblage were based on 61 consensus sequences
were obtained from 57 eimeriid coccidian species. The functional outgroup for all
phylogenetic analyses consisted of Lankesterella minima and Caryospora bigenetica.
Accession numbers for all sequences used in all analyses are included on (Appendix 4
Supplementary Table12.1).

12.4. RESULTS

12.4.1. NEWLY GENERATED SEQUENCES

All newly generated COI sequences and 18S rDNA sequences have been
submitted to GenBank. Accession numbers for COI sequences are as follow: E. ahsata
Eimeria alabamensis (MEO2015-KT184376); E. bovis (MEO2015-KT184372); E. praecox (MEO2015-KT184378); E. papillata (Chob 3-KT184377); E. tamiasciuri (MEO2015-KT184375) and L. minima (MEO2015a-KT184381; MEO2015b-KT184382). Accession numbers for 18S rDNA sequences are as follows: E. dispersa (Briston-KT184338); E. gallopavonis (USKS06-01 Type A -KT184341); E. gallopavonis (USKS06-01 Type B -KT184342); E. gallopavonis (Weybridge Type A-KT184343); E. gallopavonis (Weybridge Type B -KT184344); E. meleagridis (USAR79-01 Type A -KT184347); E. meleagridis (USAR79-01 Type B -KT184348); E. ahsata (MEO2015-KT184334); E. alabamensis (MEO2015-KT184335); E. bovis (MEO2015-KT184336); E. acervulina (NC3-KT184333); E. brunetti (Guelph 80-KT184337); E. falciformis (Chob 2A-KT184339); E. falciformis (Chob 2B-KT184340); E. maxima (M6-KT184346); E. maxima (Guelph 74-KT184345); E. necatrix (Guelph 84-KT184349); E. papillata (Chob 3-KT184350); E. praecox (MEO2015_TypeA Guelph 2012-KT184351); E. praecox (MEO2015_ Type B Guelph 2012 -KT184352); E. tamiasciuri (MEO2015-KT184353); E. tenella (USDA 80-KT184354); E. vermiformis (Chob 1-KT184355); E. zuernii (MEO2015_Guelph 2007-KT184356); Caryospora bigenetica (MEO2015a-KT184331); C. bigenetica (MEO2015b-KT184332); Isospora sp. ex Serinus canaria (MEO2015_xd24-KT184357) and L. minima (MEO2015-KT184358).

12.4.2. PHYLOGENETIC ANALYSES - CONCATENATED 18S rDNA AND COI SEQUENCES

Phylogenetic analyses based on the concatenated dataset (18S rDNA and COI) consisted of 61 single or consensus sequences from 57 coccidian species, including 4 (E. meleagrimitis, E. praecox, E. mitis and E. maxima) that were hypothesized to have two 18S rDNA paralogs in their nuclear genomes, Type A and Type B. The concatenated dataset had a total length of 2565bp. Rooted trees based on BI and ML analyses illustrated similar tree topologies (Figure 12.1, only BI trees are shown). The four Eimeria species that had Type A and Type B 18S rDNA sequences (and thus two concatenated sequences in the concatenated dataset) were each found with their corresponding Type A and Type B concatenated sequences within a monophyletic clade (PP=1.00; BP=99%). Trees based on the concatenated dataset did not support the
Figure 12.1. Phylogenetic tree based on the concatenated 18S rDNA and COI sequence dataset partitioned for Bayesian analysis as follows: CDS regions were analyzed using a codon based likelihood model (i.e. GTR+I+G (Nst=6) with Nucmodel=Codon and Code=MetMt); rDNA was analyzed simultaneously using a simpler likelihood model (i.e. GTR+I+G with Nucmodel=4by4). Missing data was treated as unknown ‘?’ and not characters for the COI partition. Monophyletic clades of Eimeria spp. were interspersed with Cyclospora or Isospora species making the genus Eimeria paraphyletic; the genus Isospora was also shown to be paraphyletic. The Eimeria species infecting turkeys are paraphyletic; E. innocua and E. dispersa formed a monophyletic clade that was only distantly related to the other Eimeria spp that infect turkeys. Eimeria species infecting chickens were also paraphyletic.
monophyly of species within either of the eimeriid genera *Eimeria* or *Isospora*. Clades of monophyletic *Eimeria* spp. were interspersed among *Cyclospora* or *Isospora* species. The *Eimeria* species infecting turkeys did not form a monophyletic group. *Eimeria adenoeides*, *E. meleagrimitis*, *E. meleagridis* and *E. gallopavonis* formed a reasonably well-supported (PP=0.97; BS=68%) monophyletic group; however, the other two species that infect turkeys, *E. innocua* and *E. dispersa*, formed a monophyletic clade (PP=1.00; BS=100%) that was only distantly related to the other *Eimeria* spp infecting turkeys. The latter two species branched basally to all other *Eimeria* species that infect galliform birds plus the human pathogen *C. cayetanensis*; thus, even the *Eimeria* species that infect galliform birds did not form a monophyletic clade.

Similarly, the *Eimeria* species infecting chickens did not form a monophyletic group. Monophyly of 5 *Eimeria* species (represented by 8 concatenated sequences) infecting chickens (*E. brunetti*, *E. acervulina*, *E. maxima*, *E. mitis*, and *E. praecox*) was well-supported (PP=1.00; BS=97%) and *Eimeria necatrix* and *E. tenella* formed a second well-supported (PP=1.00; BS=100%) monophyletic clade. The large clade of *Eimeria* species infecting chickens was the sister group to a clade of parasites of various galliform birds, including chickens (i.e. *E. tenella* and *E. necatrix*), turkeys (i.e. *E. meleagridis*, *E. adenoeides*, *E. gallopavonis* and *E. meleagrimitis*), peafowl (i.e. *E. pavonina*), and common pheasant (i.e. *Eimeria* sp. ex. *Phasianus colchicus*).

*Isospora* and *Eimeria* species that parasitize closely related definitive hosts formed well-supported monophyletic clades. For example, *Isospora* species from 3 species of passeriform birds were monophyletic (PP=1.00; BS ≤ 90%) and *Eimeria* species infecting ruminants were monophyletic (PP≤0.9; BS=100%) support. Similarly *Eimeria* species infecting rabbits (PP=1.00; BS=100%), and marsupials (PP=0.99; BS=90%) formed monophyletic clades. However, *Eimeria* species infecting rodents were not monophyletic.

12.4.3. Phylogenetic Analyses Using mt COI

The phylogenetic analysis using mt COI sequences consisted of 57 sequences from 57 coccidian species that easily assembled into an alignment, 761bp in length, with no
Figure 12.2. Phylogenetic tree based on the COI sequence dataset. Analysis was performed using a codon based likelihood model (i.e. GTR+I+G (Nst=6) with Nucmodel=Codon and Code=MetMt. Missing data were treated as unknowns, ‘?’, instead of characters. Parasites infecting closely related hosts were generally found to be closely related in the COI-based tree except for *Eimeria* spp. infecting ruminants that were paraphyletic based on COI sequences alone.
gaps. The contributing sequences ranged from 347bp (in *E. caviae* - JQ993689) to 761bp (in 29 full-length sequences). Rooted phylogenetic trees based on BI and ML analyses illustrated similar topology in the branching order for all species used in the analysis.

The overall topology of the phylogenetic trees based on the COI dataset (e.g. BI - Figure 12.2) were similar to the topology of the trees obtained using the concatenated dataset (Figure 12.1). Clades of closely related taxa (with relatively little hypothesized sequence divergence) were well supported in the COI - based trees (Figure 12.2). For example, *Isospora* spp. of passeriform birds, *Eimeria* spp of rabbits, and *Eimeria* spp of marsupials were each monophyletic with strong support (PP=1). In contrast, *Eimeria* spp. of ruminants, a group that was monophyletic with reasonable support (PP=0.90) in the concatenated dataset analysis (see Figure 12.1), were hypothesized to be paraphyletic based on COI sequences alone (Figure 12.2).

12.4.4. Phylogenetic analyses using 18S rDNA sequences.

Phylogenetic analyses based on 18S rDNA dataset also consisted of 61 single or consensus sequences from 57 coccidian species. Sequence lengths excluding gaps ranged from 902 bp (in *Isospora* sp. (JQ993670) infecting the European mole *Talpa europaea*) to 1753bp (in *Eimeria dispersa* [consensus alignment]). The general time reversible model with discrete Gamma (GTR + I + G) distribution of nucleotide substitution was determined as the best fit model. Rooted trees based on BI and ML analyses illustrated similar topologies. In all analyses based on the 18S rDNA dataset, the Type A and Type B sequences from a single parasite species were found within monophyletic clades (PP≥0.99; BS=77% - 99%).

As with all other datasets, *Eimeria* species infecting turkeys did not form a monophyletic group. Consistently, *E. dispersa* and *E. innocua* branched at the base of the large clade of *Eimeria* species infecting galliform birds and were the sister taxa to all other *Eimeria* species (plus *C. cayetanensis*) within this clade (Figure 12.3). The branching order within the clade of *Eimeria* spp. infecting turkeys (i.e. *E. adenoeides*, *E. meleagridis*, *E. gallopavonis* and *E. meleagrimitis*) was the same in all analyses. Monophyly of 5 of the *Eimeria* species infecting chickens (i.e. *E. acervulina*, *E. brunetti*, *E. mitis*, *E. praecox* and *E. maxima*) was supported in all analyses, although the
Figure 12.3. Phylogenetic tree generated from the 18S rDNA dataset. Analysis was performed using the general time reversible evolutionary model with discrete Gamma (GTR + I + G) distribution of nucleotide substitution. Species possessing the Type A and B sequences formed monophyletic clades (PP ≥ 0.99; BS = 77% - 99%). Eimeria species infecting turkeys are paraphyletic. Eimeria dispersa and E. innocua are found branching at the base of the large clade of Eimeria species infecting galliform birds. Eimeria species infecting chickens are also paraphyletic.
branching order varied slightly when analyses based on the 18S rDNA and concatenated datasets were compared to analyses based on the COI dataset.

12.5. DISCUSSION

Phylogenetic reconstruction under Bayesian and Maximum likelihood evolutionary models using 18S rDNA and partial COI sequences supported the polyphyletic origin of Eimeria species infecting turkeys. Although, the phylogenetic analyses based on BI and ML models using COI, 18S rDNA and concatenated datasets gave support to a common ancestor for many turkey coccidia, E. dispersa and E. innocua were clustered in a monophyletic clade (PP=1.00; BS 98%-99%) that was distantly related to the other species infecting turkeys. Ogedengbe et al. (2014) suggested that E. dispersa may not have evolved within turkeys as the main host, but rather infected turkeys via a host switch from another avian host such as the Bobwhite quail (Colinus virginianus, Odontophoridae). Experimental infections showed that both E. dispersa and E. innocua can parasitize both Bobwhite quail and turkeys (Phasianidae) (see Vrba and Pakandl, 2014; 2015). However, E. innocua produces 100-fold fewer oocysts within Bobwhite quail than E. dispersa suggesting that this may not be a particularly permissive host for development of this parasite (Vrba and Pakandl, 2014). Although biological and morphological similarities exist between E. dispersa and E. innocua, including the absence of a polar body in their oocysts and endogenous development within the proximal small intestine (e.g. Moore et al., 1954; Vrba and Pakandl, 2014), E. innocua was demonstrated to infect Grey partridges (Perdix perdix) readily whereas E. dispersa did not (Vrba and Pakandl, 2015). Perhaps both E. dispersa and E. innocua were originally parasites of other, perhaps even different, galliform birds and later colonized turkeys via a host switch (Ogedengbe et al., 2014; Vrba and Pakandl, 2014). The morphologically similar turkey coccidium E. subrotunda was, unfortunately, not available for study.

Phylogenetic reconstructions using each of the three datasets suggested that Eimeria species infecting chickens did not form a monophyletic group. On the basis of biological and molecular data, Barta et al. (1997) and Vrba and Pakandl (2014; 2015) suggested E. tenella and E. necatrix may have evolved from coccidia that originally
parasitized turkeys but switched to chickens during domestication of these birds. Ogedengbe et al. (2014) reported the similar observations using complete mitochondrial genome sequences for sixteen *Eimeria* species infecting chicken and turkeys. The concatenated data in the present study as well as nuclear 18S rDNA, partial mt COI sequences (Barta et al., 1997; Miska et al., 2010; Vrba and Pakandl, 2015) or complete mt genome sequences (Ogedengbe et al., 2014) all support the hypothesis for two distinct ancestral *Eimeria* sp. lineages that colonized chickens. Whether this arose from sympatric speciation of an *Eimeria* species ancestor in a primitive galliform bird followed by coevolution of hosts and parasites or by host-switching and niche specialization or some combination of the two processes, strict host-parasite coevolution between the *Eimeria* spp infecting galliform birds has likely not occurred (Vrba and Pakandl, 2015).

The 18S rDNA sequences suggested distinct paralogs in some *Eimeria* species as has been reported previously (e.g. El-Sherry et al., 2013; Vrba et al., 2011; Vrba and Pakandl, 2014). Substantial genetic divergences (2.3-2.8%) were found between apparent paralogous 18S rDNA sequences for *E. meleagrimitis* and *E. maxima*. More moderate genetic divergence was detected between paralogs of *E. mitis* and *E. praecox* 18S rDNA sequences (1.7% and 0.3%, respectively). Despite considerable sequence variation between paralogs, paralogous rDNA sequences formed monophyletic clades representing each species in the 18S rDNA and concatenated dataset analyses. Divergent 18S rDNA paralogs have been reported in other apicomplexan parasites. Qari et al. (1994) examined Type A and Type C genes from various *Plasmodium* spp. and identified 91.52% sequence similarity between these paralogs in *P. vivax*, 88.6% for *P. falciparum* and 90.17% for *P. cynomolgi*. Dame et al. (1984) reported the occurrence of four ribosomal RNA genes grouped into two major sequence classes in *P. berghei*. Stenger et al. (2015) reported highly divergent Type A and Type B 18S rDNA paralogs with sequence similarity <93% in a *Cryptosporidium* sp. infecting Eastern chipmunks (*Tamias striatus*).

*Cyclospora cayetanensis* consistently clustered with *Eimeria* spp. In previous phylogenetic analyses based on the 18S rRNA gene, this human coccidium was closely
related to *Eimeria* spp. (e.g. Relman et al., 1996; Shields and Olson, 2003). Similarly, in the present study, *C. cayetanensis* was found to closely associate with *Eimeria* spp. infecting galliforms. The close association of *C. cayetanensis* with *Eimeria* spp. prompted the suggestion by Pieniazek and Herwaldt, (1997) to reclassify this coccidium as an *Eimeria* sp. However, oocyst configuration (2 sporocysts with Stieda bodies, each containing 2 sporozoites; see Ortega et al., 1994) does not match the definition for the genus *Eimeria*, Schneider 1875. In all phylogenetic analyses, *Isospora* species infecting Passeriform birds formed a monophyletic clade intermixed among *Eimeria* species and not with *Isospora* spp. infecting rodents; thus *Isospora* spp. did not form a monophyletic clade. The genus *Eimeria* is undoubtedly paraphyletic. Repeated analyses with multiple genetic loci have concluded that the genus *Eimeria* contains a wide range of biologically diverse parasites that are held together in a single genus by virtue of their oocyst morphology and configuration. Members of the genus *Eimeria* Schneider 1875 all possess “tetrasporocystic with dizoic, nonbivalved sporocysts with or without Stieda bodies” as exemplified by the genotype, *Eimeria falciformis* (Eimer 1870) Schneider 1875 (see Upton et al, 1984; Upton, 2000). This broad generic definition has been narrowed slowly as some species have been reassigned to other genera based on oocyst configuration and life history data. For example, over 130 described *Eimeria* spp infecting fish have been reclassified into various genera (e.g. *Epieimeria*, *Goussia* or *Crystallospora*) based on the presence of valvular sutures on their sporocyst walls, location of merogony and gametogony (Dyková and Lom, 1981; Upton et al., 1984). Many parasites possessing oocysts without Stieda bodies that were previously placed within the genus *Eimeria* have been subsequently assigned to other genera. Molecular analyses (e.g. Carreno and Barta, 1999; Jirků et al., 2002) suggest that the Stieda body is a shared, derived morphological trait that is consistently expressed in sporulated oocysts of eimeriid coccidia (e.g. Carreno and Barta, 1999; Jirků et al., 2002), even when the oocysts are morphologically abnormal otherwise (e.g. Matsui et al., 1989).

Despite the morphologically obvious differences in oocyst morphology in the coccidian genera *Isospora* and *Eimeria*, monophyly of each of these oocyst morphotypes was not supported suggesting that oocyst morphology may be a poor predictor of genetic relatedness. Oocysts containing atypical numbers of functional sporocysts and
sporozoites can develop under adverse sporulation conditions. For example, *Eimeria intestinalis* oocysts exposed to elevated temperatures prior to sporulation produced a small number of *Isospora*-type oocysts containing 2 sporocysts with Stieda bodies, each containing 4 sporozoites (Matsui et al., 1989). In the same study, oocysts of the isosporoid coccidium *Cystoisospora canis* could be induced to produce *Caryospora*-like oocysts (e.g. 1 sporocyst containing 8 sporozoites) or *Tyzzeria*-like oocysts (e.g. 1 oocyst containing 8 naked sporozoites); Matsui et al. (1993) later showed that the *Caryospora*-like oocysts were as infectious as the normal oocyst morphotype, but the progeny oocysts produced during the infections initiated with the abnormal oocysts sporulated normally to typical *Cystoisospora*-like oocysts. Bisporocystic (*Isospora*-like) oocysts could be induced in *Eimeria maxima* by treating chicks with Lerbek® (a mixture of clopidol and methylbenzoquate) during infections (Norton and Joyner, 1978). The mixture of various oocyst morphotypes in the present analyses and the apparent phenotypic plasticity (Joyner 1982; Ghimire, 2010) of these oocysts suggest that how sporozoites are divided among various numbers of Stieda body possessing sporocysts may be a less consistently reliable taxonomic marker than previously assumed.

A taxonomic mess results from the demonstrated phenotypic plasticity (Joyner 1982; Ghimire, 2010) of these eimeriid coccidia; basic oocyst morphology may not be a reliable criterion at the generic level. Do we split the many described *Eimeria* species into a large number of new genera that could then each be monophyletic? Or, do we lump all coccidia with sporocysts possessing Stieda bodies into a single speciose genus, effectively synonymizing the genera *Cyclospora, Isospora, Caryospora* and *Eimeria* in the process? Neither approach is likely to be satisfactory but increasing the number of described genera is more likely to properly represent the biological diversity demonstrated by parasites currently assigned to the single genus *Eimeria*. 
13. SMALL IS NOT BORING: MITOCHONDRIAL GENOME AND SEQUENCE DIVERSITY IN THE PROTISTAN PHYLUM APICOMPLEXA

13.1. ABSTRACT

Simultaneous comparison of diverse apicomplexan mitochondrial genomes using the progressive Mauve global aligner identified conserved regions with apparent homology across multiple genome sequences. Blocks of highly similar sequences were found in varied positions among these genomes. Observed homologous local collinear blocks were detected despite numerous rearrangements among genome sequences. These were illustrated as identically coloured blocks within multiple genomes. Numerous and varied rearrangements were documented among apicomplexan mitochondrial genomes including inversions, gene translocations and variable gene content. The tissue coccidia examined (Toxoplasma, Hammondia, Neospora and Cystoisospora spp.) have a reduced number of protein-coding genes with complete COI and COIII, and a partial CytB gene (i.e. in Cystoisospora spp.) or complete COI and CytB with no evidence of a COIII gene in all other tissue coccidia examined.

Phylogenetic analyses for several apicomplexan taxa based on nucleotide sequences of concatenated mitochondrial protein coding genes and amino acids translations produced well resolved and well supported monophyletic clades for the major groups within the phylum Apicomplexa. These phylogenetic results largely support traditional or biological hypotheses of relationships within the phylum Apicomplexa, both at the species and deeper levels. Mitochondrial genome sequences appear to be useful markers for reconstructing phylogenetic relationships among both closely and distantly related taxa.

13.1. INTRODUCTION

Mitochondria are present in almost all eukaryotes as double-membrane-bound organelles functioning as the cells’ powerhouses (Henze and Martin, 2003). Mitochondrial (mt) genomes in eukaryotes are thought to have an endosymbiotic origin (Margulis and Bermudes, 1985) and believed to be derived from an alpha-proteobacterial endosymbiotic ancestor (Gray, 1999). The initial impetus that ultimately led to endosymbiosis was the need for anaerobic cells to escape from increasing
atmospheric oxygen resulting from photosynthetic activities of cyanobacteria (Andersson and Kurland, 1999). To survive, these anaerobes needed new microenvironments with less available oxygen. One solution was to permit endosymbiotic relationships with alpha proteobacteria that were capable of utilizing the excess oxygen within their intracellular environment thereby permitting these ‘anaerobes’ to persist in oxygenated environments (Andersson and Kurland, 1999). The resultant endosymbiotic relationship led to the permanent acquisition of Alphaproteobacteria via endocytosis, these permanent endosymbionts ultimately transformed into the mitochondrion (Gray, 1998; Andersson and Kurland, 1999; Thrash et al., 2011).

Mitochondrial genomes vary markedly among eukaryotes. Some are large and contain a wide variety of genes encoding many products, the so-called ‘ancestral’ mt genomes, whereas others possess fewer genes and are considered ‘derived’. The apicomplexan mt genomes are the most extreme of the “derived” types of mt genomes in eukaryotes (Gray et al., 1999; Feagin, 2000; Gray, 2012). Their mt genomes are highly reduced and some of the smallest known at only 6-11 kilobases (Feagin et al., 1991; Feagin, 2000; Gray et al., 2004; Gray, 2012). Apicomplexan mt genomes demonstrate diversity in genome structure and organization, but they generally have relatively constant complement of genes.

The physical form of apicomplexan mt genomes is variable, comprising circular or tandemly repeated copies (see Feagin et al., 1991; Wilson et al., 1992; Feagin, 1992; Preiser et al., 1996; Wilson and Williamson, 1997), arranged in a head to tail configuration (Vaidya and Arasu, 1987). Tandemly repeated mt genome forms are predominant in most Plasmodium spp. (Feagin, 2000). This form ensures that genome maps for these parasites are circular-mapping. In the piroplasmid parasites in the genera Babesia and Theileria, mt genomes occur as linear monomeric molecules with terminal inverted repeats (TIR) at both ends (Kairo et al., 1994; Omori et al., 2007; Lau, 2009; Hikosaka et al., 2011; 2012). A linear monomeric mt genome with no TIR was reported for the eimeriid coccidia Cyclospora cayetanensis (see Ogedengbe et al., 2015a). Other eimeriid coccidia in the genera Eimeria (Hikosaka et al., 2010; Lin et al., 2011; Tian et
al., 2013; Ogedengbe et al., 2013; 2014; Hafeez et al., 2015), *Caryospora* and *Isospora* (Ogedengbe and Barta, 2015; Ogedengbe et al., 2015b) and *Lankesterella minima* (Ogedengbe et al., unpublished data) possess mt genomes that are linear concatemers (circular-mapping) but there may be minor sequence variation between copies. Over 550 complete apicomplexan mt genomes representing 13 genera and 59 species have been completed, including: haemosporinids and relatives (i.e. *Plasmodium, Haemoproteus, Parahaemoproteus, Hepatocystis* and *Leucocytozoon* spp., Wilson and Williamson, 1997; Feagin et al., 2000; 2012; Perkins, 2008; Omori et al., 2008), piroplasms (i.e. *Babesia* and *Theileria* spp., Hikosaka et al., 2010; 2012), eimeriid coccidia (i.e. *Eimeria, Cyclospora, Caryospora, Isospora* or *Lankesterella* spp., Hikosaka et al., 2010; Lin et al., 2011; Liu et al., 2012; Tian et al., 2013; Ogedengbe et al., 2013; 2014; Hafeez et al., 2015; Ogedengbe and Barta, 2015; Ogedengbe et al., 2015a; 2015b; unpublished data); adeleid coccidia (i.e. *Hepatozoon* spp., Leveille et al., 2014). In these mt genomes from a wide range of apicomplexan taxa, the mt genomes contain three protein–coding genes (CDS) that encode cytochrome c oxidase subunits I (COI), cytochrome c oxidase subunit III (COIII) and cytochrome b (CytB). The protein–coding genes are interspersed among the highly fragmented and incomplete large subunit (LSU) and small subunit (SSU) ribosomal RNA genes. Within the tissue coccidia, (e.g. *Toxoplasma gondii, Hammondia heydorni* and *H. trifittae*) Gjerde, (2013) reported abbreviated mt genomes as well as shorter sequences he identified as nuclear mitochondrial DNA (NUMTs). Similar abbreviated genomes have been observed in other tissue coccidia (e.g. *Cystoisospora felis* and *C. suis*; Ogedengbe and Barta, unpublished data). In general, apicomplexan mt genomes have highly conserved gene contents and arrangements within classical groups; in the case of the eimeriid coccidia, both the CDS and rDNA fragments were highly conserved in order and primary sequences (Ogedengbe et al., 2014). Among these mt genomes of eimeriid species, sequences of the CDS regions were considerably more divergent than the rDNA fragments. The COI gene was the most conserved of the three CDS while COIII gene was most divergent (Ogedengbe et al., 2013, 2014).

Transcriptional directions of mt genes differ in apicomplexan mt genomes. If whole genome alignments were linearized at the same position (e.g. to begin at the small
subunit rDNA fragment SSU/A), translational directions for the protein and rRNA genes were oriented in the same forward direction for all eimeriid coccidia (Ogedengbe et al., 2014). In haemosporinids and their relatives (Omori et al., 2007; 2008; Hikosaka et al., 2010; Liu et al., 2010), COI and CytB genes were similarly oriented in the forward direction relative to the eimeriid coccidia, the COIII gene was however transcribed in the reverse direction to COI and CytB genes. In the adeleid *Hepatozoon catesbianae*, transcriptional directions for CytB and COIII genes were arranged in the forward orientation with respect to the direction in eimeriid coccidia, in contrast however, the COI gene was flipped both in orientation and location (Leveille et al., 2014). In the piroplasms *Babesia microti* and *B. rodhaini*, apparently continuous ‘flip-flop’ recombination events resulted in the formation of four distinct mt genome structures within a single species, each with a distinct genome map and translational directions for all three CDS (Hikosaka et al., 2012).

Most multiple alignments systems are unable to detect all evolutionary events (e.g. paralogs, orthologs, xenologous genes or lateral gene transfer) because the algorithms are assessing local (and sometimes more distant) sequence similarities when aligning multiple genome sequences (Brudno et al., 2003; Darling et al., 2004; 2007; Ovcharenko et al., 2005; Treangen et al., 2006). The Mauve aligner is a global genome alignment platform that can identify and align similar, but potentially discontinuous, regions simultaneously among a group of genome sequences (Darling et al., 2004). Mauve can help to identify some patterns of evolutionary change, particularly gene rearrangements, inversions or recombinations, in related genomes with considerable structural divergence (see Darling et al., 2004; 2007; Rissman et al., 2009).

Evolutionary studies for Apicomplexa have been based on nuclear 18S rDNA sequence analyses (Escalante and Ayala, 1995; Ellis et al., 1995; Carreno et al., 1999; Mathew et al., 2000; Perkins and Keller, 2001; Barta and Thompson, 2006; Morrison, 2009; Whipps et al., 2012). Complete mt genome sequences for over 550 Apicomplexa species are now available to be exploited in evolutionary studies. Many studies (i.e. evolutionary and ancient divergence studies and comparative gene arrangements) have included mt genome sequences in the phylogenetic analyses of wide range of taxa (e.g.
In these studies whole mt genome sequences or concatenated CDS sequences that included or excluded third codon positions, rDNA and tRNAs have been used. Many of these studies however were conducted for higher vertebrates. Similar datasets have not been investigated more extensively for apicomplexan parasites. Limited studies have focused on differences in patterns of nucleotide sequence e.g. in malaria parasite lineages (Silva et al., 2011) or use of partial CDS (i.e. COI, for species delimitation studies, Ogedengbe et al., 2011; CytB, for species identification studies, Liu et al., 2010).

This study aims to infer evolutionary diversity among apicomplexan mt genomes. The progressive Mauve global aligner will be used to simultaneously detect and identify multiple conserved genome regions. Particular effort is made to ascertain occurrence of indels, inversions, recombination, rearrangements or possible gene duplication events within the genomes of closely or distantly related Apicomplexa taxa. This study also aims to infer phylogenetic relationships between various apicomplexan parasites based on concatenated mt CDS nucleotide sequences and amino acid translations.

13.2. MATERIAL AND METHODS

13.2.1. SEQUENCES

A selection of 55 complete and 7 partial mt genome sequences from 13 genera representative of 5 major groups within the phylum Apicomplexa (i.e. 5 haemosporinid genera, 2 piroplasm genera, 5 eimeriid genera, 1 adeleid genus and 4 genera of tissue coccidia) were obtained for study; a mt genome sequence from Perkinsus marinus was obtained as a taxonomic outgroup (see Table13.1 for strains and accession numbers).

A subset of 16 complete or partial apicomplexan mt genome sequences (underlined in Table13.1) was used in the Mauve analyses. Sequences were selected to represent the various organizational forms of apicomplexan mt genomes without unnecessary duplication.

The variability in gene orders and directions precluded alignment of complete genome sequences. Therefore, phylogenetic analyses were completed using
concatenated COI, COIII and CytB sequences extracted from each whole genome. The concatenated CDS were used in subsequent phylogenetic analyses.

**Table 13.1.** Mitochondrial genome sequences used in the analyses. Underlined species were included in the Mauve analyses.

<table>
<thead>
<tr>
<th>Species</th>
<th>Stock/Strain</th>
<th>Accession Number</th>
<th>Genome Size (BP)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babesia bigemina</td>
<td>Kochinda</td>
<td>AB499085</td>
<td>5924</td>
<td>Hikosaka et al., 2010</td>
</tr>
<tr>
<td>Babesia bovis</td>
<td>Miyama</td>
<td>AB499088</td>
<td>5970</td>
<td>Hikosaka et al., 2010</td>
</tr>
<tr>
<td>Babesia caballi</td>
<td>USDA</td>
<td>AB499086</td>
<td>5847</td>
<td>Hikosaka et al., 2010</td>
</tr>
<tr>
<td>Babesia equi</td>
<td>WA</td>
<td>NW004675959</td>
<td>9001</td>
<td>Kappmeyer et al., 2012</td>
</tr>
<tr>
<td>Babesia microti</td>
<td>RI</td>
<td>FO082868</td>
<td>11,149</td>
<td>Cornillot et al., 2012</td>
</tr>
<tr>
<td>Babesia microti</td>
<td>Munich</td>
<td>AB624353</td>
<td>11,109</td>
<td>Hikosaka et al., 2012</td>
</tr>
<tr>
<td>Babesia microti</td>
<td>Munich</td>
<td>AB624354</td>
<td>11,109</td>
<td>Hikosaka et al., 2012</td>
</tr>
<tr>
<td>Babesia microti</td>
<td>Munich</td>
<td>AB624355</td>
<td>11,109</td>
<td>Hikosaka et al., 2012</td>
</tr>
<tr>
<td>Babesia microti</td>
<td>Munich</td>
<td>AB624356</td>
<td>11,109</td>
<td>Hikosaka et al., 2012</td>
</tr>
<tr>
<td>Babesia gibsoni</td>
<td>HUVIMdr1</td>
<td>AB685183</td>
<td>5867</td>
<td>Wickramasekara et al., 2012</td>
</tr>
<tr>
<td>Babesia rodhaini</td>
<td>Australian</td>
<td>AB624358</td>
<td>6929</td>
<td>Hikosaka et al., 2012</td>
</tr>
<tr>
<td>Theileria annulata</td>
<td>Ankara</td>
<td>NT_167255</td>
<td>5905</td>
<td>Pain et al., 2005</td>
</tr>
<tr>
<td>Theileria equi</td>
<td>USDA</td>
<td>AB499091</td>
<td>8246</td>
<td>Hikosaka et al., 2010</td>
</tr>
<tr>
<td>Theileria orientalis</td>
<td>Ikeda</td>
<td>AB499090</td>
<td>5957</td>
<td>Hikosaka et al., 2010</td>
</tr>
<tr>
<td>Theileria parva</td>
<td>Muguga</td>
<td>AB499089</td>
<td>5924</td>
<td>Hikosaka et al., 2010</td>
</tr>
<tr>
<td>Hepatozoon catesbianae</td>
<td>N/A</td>
<td>KF894962</td>
<td>6397</td>
<td>Leveille et al., 2015</td>
</tr>
<tr>
<td>Hepatozoon clamaeae</td>
<td>N/A</td>
<td>-</td>
<td>6377</td>
<td>Leveille et al., unpublished</td>
</tr>
<tr>
<td>Leucocytozoon caullervi</td>
<td>Gifu</td>
<td>AB302215</td>
<td>5959</td>
<td>Omori et al., 2008</td>
</tr>
<tr>
<td>Leucocytozoon fregillinarum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leucocytozoon majoris</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leucocytozoon sabrazesi</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haemoproteus columbae</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hepatoctis sp.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parahaemoproteus vireonis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasmodium vivax</td>
<td>Sal-1</td>
<td>AY598140</td>
<td>5990</td>
<td>Jongwutiwes et al., 2005</td>
</tr>
<tr>
<td>Plasmodium berghii</td>
<td></td>
<td>AB558173</td>
<td>5957</td>
<td>Nagata et al., 2010 (GB)</td>
</tr>
<tr>
<td>Plasmodium cynomolgi</td>
<td>ATCC 30146</td>
<td>AB444131</td>
<td>5984</td>
<td>Sawai et al., 2010</td>
</tr>
<tr>
<td>Plasmodium falciparum</td>
<td>ATCC 30164</td>
<td>AB444132</td>
<td>5988</td>
<td>Sawai et al., 2010</td>
</tr>
<tr>
<td>Plasmodium falciparum</td>
<td>ATCC 30164</td>
<td>AB444132</td>
<td>5988</td>
<td>Sawai et al., 2010</td>
</tr>
<tr>
<td>Plasmodium fragile</td>
<td>ATCC 30127</td>
<td>AB444112</td>
<td>5980</td>
<td>Sawai et al., 2010</td>
</tr>
<tr>
<td>Plasmodium knowlesi</td>
<td>Malaysian MRA-487</td>
<td>NC_007232</td>
<td>5957</td>
<td>Jongwutiwes et al., 2005</td>
</tr>
<tr>
<td>Plasmodium malariae</td>
<td>Oumu</td>
<td>AB489193</td>
<td>5969</td>
<td>Hayakawa et al., 2009 (GB)</td>
</tr>
<tr>
<td>Plasmodium ovale</td>
<td>AB354571</td>
<td>NC_007232</td>
<td>5974</td>
<td>Hayakawa et al., 2008</td>
</tr>
<tr>
<td>Plasmodium simium</td>
<td>Fonseca MRA-353</td>
<td>NC_007232</td>
<td>5990</td>
<td>Jongwutiwes et al., 2005</td>
</tr>
<tr>
<td>Eimeria acervulina</td>
<td></td>
<td>HJ702479</td>
<td>6179</td>
<td>Lin et al., 2011</td>
</tr>
<tr>
<td>Eimeria adenooides</td>
<td>Guelph</td>
<td>KJ608415</td>
<td>6211</td>
<td>Ogedengbe et al., 2014</td>
</tr>
<tr>
<td>Eimeria dispersa</td>
<td>Briston</td>
<td>KJ608416</td>
<td>6238</td>
<td>Ogedengbe et al., 2014</td>
</tr>
<tr>
<td>Eimeria innocua</td>
<td>European</td>
<td>KR108296</td>
<td>6415</td>
<td>Hafeez et al., 2015</td>
</tr>
<tr>
<td>Eimeria gallopavonis</td>
<td>Weybridge</td>
<td>KJ608413</td>
<td>6215</td>
<td>Ogedengbe et al., 2014</td>
</tr>
<tr>
<td>Eimeria magna</td>
<td></td>
<td>KJ608417</td>
<td>6215</td>
<td>Ogedengbe et al., 2014</td>
</tr>
<tr>
<td>Eimeria maxima</td>
<td></td>
<td>KJ608415</td>
<td>6211</td>
<td>Liu et al., 2012</td>
</tr>
<tr>
<td>Eimeria meleagrisidis</td>
<td>USA R97-01</td>
<td>KJ608418</td>
<td>6121</td>
<td>Ogedengbe et al., 2014</td>
</tr>
<tr>
<td>Eimeria meleagrinimitis</td>
<td>USMN08-01</td>
<td>KJ608414</td>
<td>6165</td>
<td>Ogedengbe et al., 2014</td>
</tr>
<tr>
<td>Eimeria mitis</td>
<td>JN864949</td>
<td>6407</td>
<td>Liu et al., 2012</td>
<td></td>
</tr>
<tr>
<td>Eimeria mitis</td>
<td>USDA 50</td>
<td>KF501573</td>
<td>6408</td>
<td>Ogedengbe et al., 2014</td>
</tr>
<tr>
<td>Eimeria necatrix</td>
<td></td>
<td>HQ702482</td>
<td>6214</td>
<td>Liu et al., 2012</td>
</tr>
<tr>
<td>Eimeria praecox</td>
<td></td>
<td>HQ702483</td>
<td>6174</td>
<td>Liu et al., 2012</td>
</tr>
<tr>
<td>Eimeria tenella</td>
<td>NIAH</td>
<td>AB564272</td>
<td>6213</td>
<td>Hikosaka et al., 2011</td>
</tr>
<tr>
<td>Eimeria tenella</td>
<td></td>
<td>HQ702484</td>
<td>6213</td>
<td>Liu et al., 2012</td>
</tr>
<tr>
<td>Carpospora bigenetrica</td>
<td>MTZ_07-12108</td>
<td>KP658102</td>
<td>6313</td>
<td>Ogedengbe &amp; Barta, 2015</td>
</tr>
<tr>
<td>Carpospora cayetanensis</td>
<td>Cyclo_CDC_2013</td>
<td>KP658101</td>
<td>6184</td>
<td>Ogedengbe et al., 2015</td>
</tr>
<tr>
<td>Isospora sp. ex. S. canaria</td>
<td>AHL_G09-038458</td>
<td>KP658103</td>
<td>6216</td>
<td>Ogedengbe et al., 2015</td>
</tr>
</tbody>
</table>
Two separate sets of alignments were constructed for the CDS to generate two datasets. For the first dataset, alignment of concatenated nucleotide sequences of CDS was performed within Geneious Version 6.1.8 bioinformatics software (http://www.geneious.com, Kearse et al., 2012) using the codon-based ‘Translation Alignment’ algorithm using the mold/protozoan mitochondrial codon translation (i.e. translation table 4). Alignment parameters were: Cost Matrix = Blossum 62 , Gap open penalty = 12 , and, Gap extension penalty = 3. Positions for which sequence data was not available (missing) were treated as unknown and were coded as such (i.e. '?' inserted). For the second dataset, the concatenated CDS nucleotide sequences were translated to amino acids; alignment was constructed on the concatenated amino acid translations and performed within Geneious using the ‘MAFFT Alignment’ algorithm (Katoh and Toh, 2010). Parameters were: Algorithm = GINS-1 , Cost Matrix = Blossum 62 , Gap open penalty = 1.53 , and, offset value = 0.123.

13.2.3. PHYLOGENETIC ANALYSES

Phylogenetic analyses were performed on the two datasets: 1) concatenated and aligned nucleotide sequences of COI, COIII and CytB and 2) concatenated amino acid translations for COI, COIII and CytB genes excluding stops.

Nucleotide-based Analysis. The nucleotide-based analysis consisted of an alignment of 63 taxa with a total length of 3561bp for the concatenated CDS. Optimal model selection was performed within the MrModel v2.3 software by Nylander J. A. A. (2004). MrModeltest v2. Program; distributed by the author, Evolutionary Biology Centre, Uppsala University. The best fit model with the lowest Akaike information criterion (AIC) score for the nucleotide sequences dataset was the general time
reversible (GTR) model (Tavaré, 1986) with invariant gamma distribution of nucleotide substitution (I+G) and 6 rate categories of nucleotide substitution. Preliminary alignments of individual CDS (i.e. COI, COIII or CytB) demonstrated the same best fit model individually so the final model selection was made on the concatenated nucleotide dataset. Bayesian Inference (BI) was executed from within MrBayes bioinformatics software Version 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). The codon nucleotide model (i.e. Nucmodel = Codon) using the mammalian mitochondrial translation code (i.e. Code = metmt) was used in the BI analysis. Missing data and gaps were treated conventionally (i.e. ‘gaps = -‘, ‘missing = ?’). A Markov Chain Monte Carlo search was set to $1 \times 10^6$ generations. The analysis was run in four chains; the heated chain default temperature was 0.2 and burn-in length at 400,000. Subsample frequency of 1,000 for all datasets. Convergence of trees was saved and examined after 1000 burnin.

Amino Acid-Based Analysis. 2) The resulting amino acid alignment was 1183 in length with 577 phylogenetic informative positions in the final dataset. The best fit evolutionary model test and selection for the amino acid translation dataset was performed in MEGA version 6.0 (Tamura et al., 2013). The LG+G+I+F matrix (Le and Gascuel, 2008) was selected as the best fit model. The LG matrix incorporates discrete gamma distribution (+G, 0.99) estimation of invariant (+I, 0.13) while accounting for variability of rates across sites plus implementing the empirical amino acids frequencies (+F, estimated as 1). Number of nucleotide substitution rate categories = 5. Phylogenetic analysis was run using the PhyML 3.0 algorithm phylogenetic estimation (Guindon et al., 2010) using Maximum Likelihood (ML) method (Guindon and Gascuel, 2003) and executed within Geneious bioinformatics software (Kearse et al., 2012) with 500 bootstrap replicates. The tree topology search heuristic ‘Best’ was selected. ‘Best’ combines the Nearest Neighbor Interchange (NNI) and Subtree Pruning and Regrafting (SPR) tree topology search options that reduces the number of topologies searched while improving the likelihood of best tree produced. Sequences extracted from the mt genome of the dinoflagellate Perkinsus marinus ATCC 50983 was used to root all resulting phylogenetic trees.
13.2.4. MAUVE ANALYSIS

The progressive Mauve analysis dataset consisted of complete mt genome sequences from 26 apicomplexan taxa (underlined in Table 13.1). These genome sequences were used in various subsets for Mauve analyses. Analysis using the progressive Mauve version 2.0 and later (see Darling et al., 2004; 2007; 2010) was executed from within Geneious bioinformatics software (Kearse et al., 2012) with the following parameters: Match seed weight = 10, minimum LCB scores = 28, compute locally collinear blocks, full alignment, extend LCBs, and MUSCLE 3.6 gapped aligner.

13.3. RESULTS

13.3.1. MAUVE ANALYSIS

Mauve result outputs presented as visuals demonstrating the extent of conservation among genes, genome rearrangements and inversions was illustrated in Figures 13.1-13.4.

Figure 13.1 demonstrated the comparative analysis on genome sequences from 16 apicomplexan parasites generated through global sequence alignments using Mauve aligner. Significant large ‘backbones’ with similar genomic contents (LCBs) were observed as conserved segments and represented in colored blocks.

Protein coding genes were individually depicted in identical colored LCB segments across the genomes (e.g. COI, green, COIII, blue and CytB, purple color); the LCB regions for COI was mostly conserved and represented among all the genome sequences. No LCB was identified for COIII in tissue coccidia. The partial sequences used in this analysis did not include the COIII gene. The degree of sequence similarity was indicated in the peak of the coloured block similarity profile. Dissimilar regions in the LCB fragments were illustrated as lesser peaks. Unaligned regions representing lineage specific sequences indicated as white spaces were predominant within the COIII LCB fragments across all the taxa.

Transcriptional directions for coding genes were diverse but conserved within classical groups, except the piroplasms. Inversions (i.e. DNA located in the same region
but in the opposite orientation) were shown in different species as collinear blocks below the genomes’ center line. Within the piroplasms, in reference to *B. microti* (Figure 13.2), COI was in forward direction in all species; CytB was inverted in all except in *B. microti*, *B. rodhaini* and *T. equi*. COIII gene was inverted in all species except in *B. microti* spp. and *T. orientalis*. This rearrangement was different from what was observed in the mt genome sequences for *Plasmodium* or *Eimeria* species. Putative translational direction for the protein and rDNA LCBs were oriented in the same direction among all *Eimeria* spp. mt genomes and among all haemosporinid parasites although directions differed between these groups of parasites. The adeleid coccidia, *Hepatozoon catesbiana*, COIII LCB was not illustrated. COIII was inverted in *Leucocytozoon*. Most inversions were however found in ribosomal genes across the genomes (Figure 13.1).

Gene rearrangements were illustrated by the array of LCB connecting lines across the genome sequences. Figures 13.2, 13.3 and 13.4a-3c illustrated heightened gene rearrangements between five piroplasms species, rearrangements and gene orders within a single piroplasm species, *B. microti*, and pairwise comparisons constructs between *Eimeria* and *Leucocytozoon* or *Hepatozoon* or *Babesia* spp. However, gene rearrangements were conserved among groups. Gene duplication was not detected.
Figure 13.1. Illustrates a screenshot of Mauve alignment for 16 mt genomes of various apicomplexan parasites, illustrating global genome rearrangements. Genomes are arranged horizontally. Locally collinear blocks representing regions of high sequence similarity are shown as coloured blocks. The elevation in the similarity contour corresponds to the degree of similarity between genome sequences. Homologous LCBs were connected with arrays of connecting lines across the genome sequences. LCB blocks that were shifted downwards below the genome’s center line represent genes inverted relative to the reference genome (P. falciparum). MAUVE parameters as follow: Match seed weight = 10; minimum LCB scores = 28; compute locally collinear blocks; full alignment; extend LCBs; and, MUSCLE 3.6 gapped aligner.
Figure 13.2. Schematic illustration of gene orders, rearrangements and inversions in four piroplasms *B. microti*, *B. rodhaini*, *T. equi* and *T. orientalis*. Regions of the same colour connected by lines are hypothesized to be homologous regions based on sequence similarities.

Figure 13.3. Variable rearrangements and gene orders within a single piroplasm species, *Babesia microti*. Upper coloured regions of each genome are encoded in the forward direction and lower coloured regions are encoded in the reverse direction.
Figures 13.4. Illustrate pairwise comparisons of gene rearrangements between the mitochondrial genome of *Eimeria mitis* and mt genomes of: a) *Leucocytozoon caulleryi*; b) *Hepatozoon catesbiana* and *Babesia rodhaini*.

13.3.2. PHYLOGENY

*Phylogenetic analyses using DNA sequences of concatenated CDS*

Phylogenetic analyses based on the concatenated CDS dataset consisted of 63 sequences with total length of 3561bp. The BI tree (Figure 13.5) supported the monophyly of the major groups within the Apicomplexa that were analysed in this study. Monophyly of the eimeriid coccidia (i.e. *Eimeria, Isospora, Caryospora, Cyclospora* and *Lankesterella* spp.), was well-supported (PP=0.99), the tissue coccidia (*Toxoplasma, Neospora, Hammondia* and *Cystoisospora* spp.) and the piroplasms (i.e. *Babesia* and *Theileria* spp.) each had a PP=1.00. The haemosporinids (i.e. *Plasmodium*, *Theileria*, *Babesia*, *Lankesterella*).
Parahaemoproteus, Leucocytozoon, Haemoproteus and Hepatocystis) and adeleid coccidia (i.e. Hepatozoon spp.) also formed well-supported monophyletic clades (PP=1.00, respectively).

Within the large clade of eimeriid taxa, the monophyly of the genus Eimeria was not supported. Eimeria species were found within a well-supported clade that also contained Isospora and Cyclospora spp. A well-supported (PP=0.98) monophyletic clade of facultatively or obligately heteroxenous eimeriid coccidia (i.e. Caryospora bigenetica and Lankesterella minima) was the sister group to this homoxenous eimeriid clade. Eimeria spp. infecting turkeys and chickens did not form monophyletic groups. Several species infecting turkeys (i.e. Eimeria adenoeides, E. meleagrimitis, E. meleagridis and E. gallopavonis) formed a well-supported (PP=1.00) monophyletic group that included two chicken species (i.e. E. tenella and E. necatrix, PP=1.00). Two other species that infect turkeys, E. dispersa and E. innocua, also formed a monophyletic clade (PP=0.99), but this clade was only distantly related to the other Eimeria spp. infecting turkeys. Five Eimeria species infecting chickens (i.e. E. acervulina, E. brunetti, E. maxima, E. mitis, and E. praecox) formed a well-supported (PP=1.00) monophyletic clade.

The tissue coccidia (Sarcocystidae) formed a well-supported (PP=1) sister clade to the eimeriid coccidia. Genera within the tissue coccidia also formed well-supported clades that were sister clades to each other (i.e. Hammondia spp. were sister clade to Neospora, PP=0.59); Toxoplasma was sister clade to the Hammondia/Neospora clade (PP=1.00), and Cystoisospora spp. formed a well-supported (PP=1.00) monophyletic clade that was the sister group to the Toxoplasma/Hammondia/Neospora group.

The haemosporinids (Plasmodium, Leucocytozoon, Hepatocystis, Haemoproteus and Parahaemoproteus spp.) formed a well-supported (PP=1.00) monophyletic clade but the monophyly of the individual genera was not supported. The Hepatocystis sp. concatenated sequence grouped within a large clade of Plasmodium spp. Likewise, Leucocytozoon caulleryi did not group with the other Leucocytozoon spp. included in the analysis. The piroplasms also formed a well-supported monophyletic clade (PP=1.00) for Babesia spp. interspersed with a well-supported (PP=1.00) clade of Theileria spp.
Figure 13.5. Phylogenetic tree based on concatenated CDS dataset. Monophyly of major groups within Apicomplexa is supported (PP ~ 1.00). The genus *Eimeria* is polyphyletic; *Eimeria* species infecting turkeys or chickens each did not form monophyletic clades. The genus *Eimeria* is interspersed with *Isospora* and *Cyclospora* species. The haemosporinid *Leucocytozoon caulleryi* did not cluster with other *Leucocytozoon* spp. The adeleid coccidia *Hepatozoon catesbianae* and *H. clamatae* branched basally to the haemosporinid parasites that share invertebrates as definitive hosts.
Phylogenetic analyses using concatenated amino acid (aa) translations

Phylogenetic analyses based on concatenated amino acid translations of the same CDS generated an alignment of 1183 (aa) translations in length derived from 63 contributing parasites complete or partial mt genome sequences. The ML tree (Figure 13.6) based on the concatenated aa translations produced similar topology with minimal modification to branching order relative to the concatenated CDS nucleotide BI based tree (Figure 13.5). Major groups comprising closely related taxa were highly supported. For example, 100% bootstrap support (BS) supported the monophyly of the major groups (i.e. eimeriid coccidia, tissue coccidia, haemosporinids, adeleids and piroplasms). ML analysis using aa translations placed the well-supported (BS =100%) monophyletic clade of the adeleids *Hepatozoon catesbianae* and *H. clamatae* branching basally to the eimeriid and tissue coccidia group compared to the concatenated coding protein nucleotide based BI tree, where these adeleid were positioned basally to the haemosporinids cluster. *Plasmodium* spp. formed a polytomy with other haemosporinids that resolved into two clades consisting of a well-supported clade of *Plasmodium* and *Hepatocystis* spp. (BS=99%) and a second weakly-supported clade of *Plasmodium, Leucocytozoon, Haemoproteus* and *Parahaemoproteus* spp. The two clades however formed a well-supported (BS =100%) larger clade for all haemosporinids coccidia.

Phylogenetic relationships of *Eimeria* species infecting turkeys and chickens and other coccidia species were identical to the relationships deduced using the concatenated protein coding nucleotide sequence dataset (Figure 13.5). The relationships between the tissue coccidia was however slightly altered in the tree based on aa translations; in this group, *Hammondia* spp. were sister clade to *Toxoplasma*, BP=90%). *Neospora* was sister clade to the *Hammondia/Toxoplasma* clade (BP=100) while the well-supported (BP=100%) *Cystoisospora* spp. formed a monophyletic clade sister group to the *Hammondia/Toxoplasma/Neospora* group.
Figure 13.6. Phylogenetic tree based on concatenated amino acid translation dataset. Major groups within Apicomplexa is supported (BP=100%). The adeleid coccidia *Hepatozoon catesbiana* and *H. clamatae* branched basally to other coccidia. Haemosporinid genera were not found to be monophyletic. The amino acid translation based tree also produced a polyphyletic clade for the genus *Eimeria*.
13.4. DISCUSSION

The current availability of more than 550 complete apicomplexan mt genomes representing 59 distinct species (>500 haemosporinid sequences from 35 species; 13 piroplasm sequences from 7 species; 23 coccidian sequences from 16 species and 1 adeleid sequence) provides a powerful tool for understanding evolutionary changes within the mt genomes of members of the Apicomplexa phylum viz a viz indels, insertions, duplications, inversions, recombination, rearrangements or translocations. Generally, protein coding and rRNA-coding gene regions were more conserved than non-coding intergenic regions (Dubchak et al., 2000; Brudno et al., 2003) due to functional constraints. It was evident from the progressive Mauve (Daring et al., 2010) analyses that changes in the transcriptional direction are common among apicomplexan parasites. Despite the tremendous variation in genome organization observed in these apicomplexan parasites; mt gene content, order and directions were largely conserved within each of the major apicomplexan groups with the exception of the piroplasms. The latter group is known to possess a bewildering array of genome structures (e.g. Cornillot et al., 2012; Hikosaka et al., 2010; 2012; Kappmeyer et al., 2012).

Transcriptional variation was most pronounced in the piroplasms. The monomeric mt genomes of Babesia equi have long terminal inverted repeats (TIR, i.e. IR-A and IR-B palindromes) at each end of their linear genomes (Nosek and Tomaska, 2003; Hikosaka et al., 2012). In other Babesia spp. the presence of TIRs mediated an apparently stable dual ‘flip-flop’ recombination phenomenon resulting in different transcriptional directions for genes located within the TIR (Hikosaka et al., 2012). Plasmodium berghei, the rodent malaria parasite is known to alter its energy metabolism from glycolysis in the mouse to oxidative phosphorylation in the mosquito definitive host as an adaptation to its heteroxenous life cycle (Hino et al., 2012). Hikosaka et al. (2012) suggested that the flip-flop recombination phenomenon in the mt genomes of B. microti and B. rodhaini (close relatives to Plasmodium) may affect translation of the CDS (similar to some recombination mechanisms of expression control used by bacteria, e.g. Cui et al., 2012b) and, perhaps, be an adaptation for living within different hosts.
Mauve analyses of apicomplexan mt genomes reflected genomes that have undergone massive rearrangement (Omori et al., 2007; 2008). Gene content, order and orientation may have resulted from mutational forces serving as precursors of evolutionary events leading to changes in genes rearrangements (Darling et al., 2007). Such mutation forces may relate to the generation of reactive oxygen species ROS as by-products of oxidative phosphorylation responsible for oxidative damage (Cui et al., 2012a) or from the subsequent recombination events during DNA repair process from oxidative damage (Bogenhagen, 1999). There is substantive evidence that mt genomes in Apicomplexa have undergone significant horizontal genes transfer (e.g. in tissue coccidia, Gjerde, 2013) as well as genome rearrangements (Omori et al., 2007; 2008; Hikosaka et al., 2010; Leveille et al., 2014). Horizontal gene transfers result in new genome contents (Darling et al., 2007). The occurrence of horizontal gene transfers (e.g. in Plasmodium, Kishore et al., 2013) or gene duplication and alternate genome structures within mt genomes (e.g. in Babesia, Hikosaka et al., 2010) may lead to novel genes and changes in substitutions of variant forms within genomes (Darling et al., 2007). Mitochondrial genes are maternally derived with limited genetic recombination. Mitochondrial genes are subjected to high rate of base substitution; these genes however lack introns that could make sequences less reliable (Curole and Kocher, 1999). The uniparental inheritance of mt genes ensure that variants (possessing synapomorphic traits) are retained and set in populations (Curole and Kocher, 1999). These factors plus the availability of a growing catalogue of mt genome sequences from biologically divergent apicomplexan parasites make mt gene sequences useful for phylogenetic analysis.

Evolutionary relationships among Eimeria species infecting chickens and turkeys using complete mt genome sequences or concatenated CDS nucleotide sequences or CDS translations data have been demonstrated (Ogedengbe et al., 2014). In the present study, concatenation of nucleotide sequences of all three CDS, amino acid data and partials were exploited in phylogenetic analysis of diverse apicomplexan taxa. Whole mt genome amplifications of the tissue coccidia have demonstrated only partial gene recovery, mainly the COI and CytB for Toxoplasma, Neospora and Hammondia spp. (Gjerde, 2013), COI and COIII for Cystoisospora spp. (Ogedengbe and Barta,
unpublished data). Criticisms have been expressed in the use of mt genomes sequences to resolve evolutionary relationships at deeper level. Curole and Kocher, (1999) described the challenge in using mt genomes for phylogenetic analyses; contradictions were encountered in the evolutionary studies involving the vertebrate coelacanth \textit{(Latimeria chalumnae)} and close relatives lungfish \textit{(Protopterus dolloi)} and tetrapods based on complete mt genome sequences and nuclear 28S rDNA sequences. Competing hypotheses were strongly supported by alternate mt genome data sets used in the phylogenetic analyses for these vertebrates. The non conventional relationships (e.g. lungfish and tetrapod clade) inferred with the quickly evolving mt genes produced well-supported topologies as did the historically (paleontological) proven accepted coelacanth and lungfish relationships, inferred with the slowly evolving nuclear genes (see Curole and Kocher, 1999). The expectation was that if mt genome sequences have limited resolving power, the non conventional topologies should have low nodal supports. In Curole and Kocher, (1999) study, the non conventional relationships continued to be the best supported hypothesis despite applying several phylogenetic models and variety of algorithms to the datasets. The question then is whether the non conventional relationships were due to Long Branch attraction (Curole and Kocher, 1999) caused by short deep divergences (Townsend et al., 2012) between these sarcopterygian lineages as the result of their rapid evolution (Zardoya et al., 1998) or due to erroneous interpretation of morphological data (Curole and Kocher, 1999).

Phylogenetic analyses for 63 apicomplexan mt genome sequences based on concatenation of CDS nucleotide sequences and aa translations generated similar tree topologies in which major clades of eimeriid and tissue coccidia, haemosporinids and related species, piroplasms and adeleids were well-supported, even when partial mt genome sequences were used for tissue coccidia. Two protein coding sequences, COI and CytB only were obtained for \textit{Toxoplasma gondii}, \textit{Hammondia heydorni}, \textit{H. triffittae} and \textit{Neospora caninum} (see Gjerde, 2013), while complete COI and COIII as well as a partial CytB CDS were obtained for \textit{Cystoisospora felis} and \textit{C. suis} by Ogedengbe and Barta (unpublished). Interestingly, phylogenetic results obtained from the use of partial CDS in the concatenated protein coding nucleotide BI tree or in the aa translation-based ML tree did not contradict the broadly accepted traditional relationships for the tissue
coccidia. All tissue coccidia (*Toxoplasma, Hammondia, Neospora* and *Cystoisospora* spp.) formed well-supported (PP=1.00; BS=100%) monophyletic clades and similarly well-supported (PP=1.00; BS=100%) monophyletic sister clade to all other eimeriid coccidia. In many analyses of tissue coccidia, the canid-infecting parasites in the genera *Hammondia* and *Neospora* form a monophyletic clade that is the sister group to a clade containing *T. gondii* and *H. hammondi* (see Chapter 11.4; Ellis et al., 1999; Mugridge et al., 1999; Mugridge et al., 2000; Morrison, 2009). This could not be tested in the present work because a mt genome sequence for *H. hammondi* was not available to include in this study. In the aa-based tree, however, the relationships among *Hammondia*, *Toxoplasma* and *Neospora* spp. did not reflect any such definitive host associations.

Phylogenetic relationships inferred among malarial parasites (*sensu latu*, Perez-Tris et al., 2005) belonging to 5 haemosporinid genera did not support the monophyly of two genera: *Leucocytozoon* and *Plasmodium* (Perkins, 2008). *Leucocytozoon* (*Akiba caulleryi*) (the only described member of the subgenus *Akiba*) uses members of the genus *Culicoides* (biting midges, family Ceratopogonidae) as vectors and did not group with other *Leucocytozoon* spp. that use members of the genus *Simulium* (blackflies) as vectors. Rather *Leucocytozoon* (*Akiba caulleryi*) associated closely with another haemosporinid parasite that also uses *Culicoides* as vectors, *Parahaemoproteus vireonis*, although these two parasites did not form a monophyletic group. The results also confirmed the paraphyly of the genus *Plasmodium* (Escalante et al., 1998; Perkins and Schall, 2002; Martinsen et al., 2008); several *Plasmodium* lineages formed a polytomy with four haemosporinid genera: *Hepatocystis, Haemoproteus, Parahaemoproteus* and *Leucocytozoon* in which associations generally conformed to either the mammalian host group (*Plasmodium* and *Hepatocystis*) and the avian/reptilian hosts group (*Plasmodium, Haemoproteus, Parahaemoproteus*, and *Leucocytozoon*) (See Rich and Xu, 2011). The four haemosporinid genera that clustered with *Plasmodium* utilize different dipteran vectors reflecting their origin from an ancestral avian parasite through vector switches between Simuliidae (Blackflies), Hippobosidae (louse flies); Ceratopogonidae (biting midges) and Culicidae and Anopheles (mosquitoes) (see Martinsen et al., 2008).
Generally, in both phylogenetic analyses, the inferred relationships were largely consistent with already established concepts (Kuo et al., 2008) of apicomplexan evolution. The traditionally recognized major groups within the Apicomplexa were all well-supported within monophyletic groups. In addition, the close relationship between the eimeriid and tissue coccidia was supported with these large assemblages of species forming sister clades. In the aa-based phylogenetic reconstruction, the adeleid coccidia formed the sister group to the other coccidia, together forming a monophyletic group comparable to the apicomplexan order Eucoccidiorida. In neither analysis was the monophyly of the Aconoindasida supported; the piroplasms and haemosporinids never formed a monophyletic group despite the strong support for such a relationship based on a large number of aa-based analyses of nuclear genes from apicomplexan parasites for which whole nuclear genomes were available (Kuo et al., 2008). In the latter paper, data from a representative of the adeleid coccidia were not available for analyses.

13.5. CONCLUSION

From this study, the progressive Mauve aligner allowed multiple comparisons of whole genome sequences. Locally collinear blocks easily detected genome regions with high similarity and illustrated the results of evolutionary events such as genome rearrangements and inversions. Conserved regions irrespective of their location or orientation were identified as multi-MUMs. Divergent regions in these genomes were less adequately aligned by Mauve. The unaligned regions can be recognized as heterologous regions specific to different lineages or just simply repetitive sequences that Mauve could not detect. Although inversions were observed, other evolutionary events such as gene duplication as previously reported in *T. equi* (Hikosaka et al., 2010) was not illustrated in the Mauve analysis. The positions and orientations of LCBs detected in this study could further be applied in studies devoted to genome rearrangements and permutations in mt gene order (Bourque and Pevzner, 2002; Larget et al., 2002; Darling et al., 2004). Such observations may also be applied to evolutionary studies for Apicomplexa taxa.

Topologies describing evolutionary relationships for all the major groups of Apicomplexa were consistent with traditional concepts of relationships deduced from
established hypothesis (i.e. ultrastructural and biological characters; see Levine et al., 1988) for these apicomplexan groups. In all analyses, closely related species clustered together and formed well-supported clades and sister groups in both the protein coding nucleotide and aa based trees. In this study evolutionary relationships based on mt genome sequence datasets resolved more relationships than expected in comparison to nuclear sequence dataset even at deep levels for the phylum Apicomplexa.

14. GENERAL DISCUSSION AND HIGHLIGHTS
14.1. GENERAL DISCUSSION

The phylum Apicomplexa is home to an incredibly speciose, most yet to be described, collection of obligately parasitic protists that include in their numbers important parasites of medical and veterinary importance. Infections by apicomplexan parasites limit the profitability or sustainability of intensive rearing of livestock and poultry and the general well-being of companion animals and humans. Research on the biology of apicomplexan parasites over many decades has generated considerable information on the life cycles, morphology or morphometrics of exogenous (i.e. oocysts, sporocysts and sporozoites) and endogenous stages (i.e. meronts, merozoites, gamonts and gametes), methods of identification and characterization, pathogenicity and disease impact on animal and human health, and evolutionary relationships among apicomplexan parasites. An important challenge in characterizing members of the phylum Apicomplexa has been differentiating closely related species. Accurate identification is an important step towards characterization, effective diagnosis and control as well as the elucidation of phylogenetic relationships among members of the phylum. Identification and characterization based on oocyst morphology and other biological traits (e.g. predilection sites, lesion formation, intermediate or definitive hosts, see Levine, 1988) alone are sometimes inadequate to differentiate species. For example, the oocysts of *Isospora greineri* and *Isospora superbusi*, two species described from Superb Glossy Starlings (*Lamprotornis superbus*) cannot be differentiated morphologically (Hafeez et al., 2014). Among the *Eimeria* species infecting chickens and turkeys, the morphological features of exogenous parasite stages (i.e. oocyst dimensions, shape and color) are consistent within a species but these same features can
overlap considerably between distinct *Eimeria* species, even within a single host (El-Sherry et al., 2015). Effective control measures may require species level identification for the formulation of live coccidiosis vaccines. In the present work (Chapter 11), morphological identification of *Cystoisospora* spp. for which oocysts were available was further complicated by the wide range of oocyst measurements reported previously (e.g. Dubey and Frenkel, 1972a; Agrawal et al., 1981; Bjork et al., 2000).

Similarly, the use of limited molecular targets, particularly nuclear 18S rDNA, has proven to be inadequate to delimit closely related species; 18S rDNA sequences did not resolve the monophyly of *Eimeria tenella* (Ogedengbe et al., 2011). Further, in many phylogenetic and related studies completed to date, data from only a limited number of taxa were available for inclusion; only a few (e.g. Whipps et al., 2012) have sufficiently robust taxon sampling to promote stability in the resulting phylogenetic hypothesis. Results from many studies suggest that numerous and substantial taxonomic changes may be required in the Apicomplexa if monophyly of genera and higher taxonomic groups within the phylum is to be maintained (Morrison et al., 2004).

Molecular identification can provide additional diagnostic tools for accurate identification of apicomplexan parasites (Chapman et al., 2013). A good genetic marker should be one with sufficient sequence variation between species. Until recently, the nuclear (nu) 18S rDNA loci have been used widely for molecular characterization and phylogenetics of various coccidia (e.g. Barta et al., 1997; Carreno and Barta, 1999; Morrison et al., 2004; Morrison, 2009). This gene however has limited use for inferring relationships among closely related species (Barta et al., 1997; Li et al., 1997; Heger et al., 2011; Ogedengbe et al., 2011). The presence of multiple paralogous copies and indels (McCutchan et al., 1988; El-Sherry et al., 2013) within a single parasite or distantly related coccidia makes the nuclear 18S rDNA sequences inadequate for species identification of the closely related species. Alignments of 18S rDNA sequences can be difficult making the assignment of positional homology almost impossible. The internal transcribed spacer (ITS) regions similarly exhibit variation in sequences within the intergenic regions that sometimes result in paralogous copies that make positional homology of sequences difficult (Alverson and Kolnick, 2005; Morgan et al., 2009).
These challenges limit the use of the nuclear genes in phylogenetic relationships of closely related species (Lew et al., 2003).

Alternatively, Hebert et al. (2003a) suggested ‘DNA Barcoding’ (the use of sequence diversity in a short 500-800bp mitochondrial cytochrome c oxidase subunit I [mt COI] gene) as a suitable marker for identification or delimitation of closely related species. The mt COI gene in protistan organisms has a sequence divergence rate that may be 1% to greater than 2% per million years (Hebert et al., 2003b; Zhao et al., 2012). The mt COI gene is maternally derived, conserved and well characterized target within taxa in which it is found. The uniparental inheritance of mt genes ensure that complications associated with recombination are reduced and therefore alleles are retained in populations (Curolle and Kocher, 1999). This quickly evolving genetic locus shows considerable interspecific sequence diversity among closely related species and comparatively little intraspecific sequence variation. This makes mt COI sequences highly suited for identification and differentiation of closely related eimeriid coccidia compared to complete or near-complete nuclear 18S rDNA sequences (Ogedengbe et al., 2011). The COI target can be used to advantage for molecular phylogeny of apicomplexan parasites. Although DNA Barcoding has been exploited on a large scale to delineate many animal species (e.g. iBOL), it has been used only infrequently for apicomplexan parasites because of lack of suitable PCR primers. Identifying parasites with COI requires a library of sequences from well-characterized parasites for comparison with sequences from unknown samples. Establishing a ‘best practices’ method of reliably identifying eimeriid and other coccidia required a series of studies (Chapters 11 and 12) that were each dependent on the availability of well-characterized, single-species samples (see Appendices 7, 8, 9, 10, 11, 12, 15.).

Progress was made in this study to identify a suitable barcode region (or regions) for the Apicomplexa, preferably within the region of the mt COI gene that overlaps with COI-based DNA barcoding initiatives with other taxonomic groups. A single primer pair was not applicable to all taxa within the phylum, so group-specific primer pairs were designed; one pair of degenerate primers was capable of amplifying a portion of the mt COI from all eimeriid coccidia tested to date (see Chapter 12) and a second primer set
was capable of amplifying a portion of the mt COI from all sarcocystid coccidia (see Chapter 11). Collectively, these primer sets can amplify DNA from virtually all monoxenous and heteroxenous coccidia.

Mitochondrial COI was identified as an effective barcode gene (Hebert and Gregory, 2005) and its use with Apicomplexa that possess a mt genome was just beginning at the start of the thesis research reported herein. This study generated necessary background information, suitable PCR primers and reference COI sequences from a taxonomically broad range of biologically characterized apicomplexan parasites to permit partial mt COI sequences be used successfully as a molecular species identification tool (DNA barcode) for many apicomplexan parasites (and most coccidia) of veterinary and medical importance (see Chapters 11-12). The optimized molecular methods and reference COI sequences that were generated should help usher in a new, broadly applicable, molecular identification method for these medically and economically important apicomplexan parasites infecting humans, domestic and wild animals globally. The specificity of the COI primers permits amplification of partial mt COI sequences from within infected tissues or fecal samples. These sequences can then be used to molecularly link different life cycle stages of a single parasite species in various tissues or hosts.

For many coccidia, mt COI-based DNA barcoding did an excellent job of differentiating even closely related parasites (Chapters 11-12). However, even DNA barcoding using the mt COI locus were unable to differentiate morphologically similar parasites in the genus *Cystoisospora*. Canid-infecting species belonging to the *Cystoisospora ohioensis* species complex (i.e. *C. ohioensis* Dubey 1975a, *C. neorivolta* Dubey and Mahrt 1978 and *C. burrowsi* Trayser and Todd 1978) produce oocysts that are morphometrically similar. Morphological and molecular assignment of representatives of the *Cystoisospora ohioensis* species complex to a particular species was difficult. COI sequences were shown to be well-suited for species delimitation of eimeriid coccidia (Ogedengbe et al., 2011) and for deducing their evolutionary relationships (Ogedengbe et al., 2011; Chapter 11). Nonetheless, partial mt COI sequences from *Cystoisospora* spp. demonstrated remarkably limited sequence variation.
For example, as examined in detail in Chapter 11, the partial COI sequence of C. cf. *ohioensis* (isolate Sedona) differed from the COI sequences of *C. suis* and *C. rivolta* by a single nucleotide difference (SND) over a 475 bp portion while differing from a second *C. cf. ohioensis* (IDX005) isolate by three SNDs. The pairwise differences in the COI sequences between two *C. cf. ohioensis* isolates may represent species-level differences within the *C. ohioensis* species complex. This suggestion can only be confirmed through generation of pure lines of each parasite and observations on the endogenous development of these pure lines to assign specific COI sequences to *C. ohioensis*, *C. burrowsi* or *C. neorivolta*. The limited sequence divergence at the mt COI locus suggests recent radiation of these parasites into a number of mammalian hosts. Failure to cross-infect dogs with feline parasites and vice versa support *C. felis* and *C. canis* as distinct species (Dubey, 1975a) despite having only 5-6 SNDs over 682bp in the COI locus, as well as *C. rivolta* and *C. ohioensis* as distinct species (Dubey, 1975b) despite having only a single SND between that latter two species. Clearly, a different DNA barcode locus (perhaps the rapidly evolving nu ribosomal ITS regions) will be required to reliably differentiate *Cystoisospora* species; however, mt COI was highly effective at assigning a particular coccidium to the genus *Cystoisospora* rather than some other genus of tissue coccidium such as *Hammondia*, *Neospora* or *Toxoplasma* (see Chapter 11).

Mitochondrial COI-based DNA barcoding readily differentiated the seven *Eimeria* species that infect chickens (Ogedengbe et al., 2011). An attempt was made to exploit this capability using SYBR green real time PCR (qPCR). Initially, qPCR was used to optimize a genus-specific detection assay for *Eimeria* species that infect chickens. Using genus-specific primers, ~200 bp products were amplified from any and all *Eimeria* species in a sample and the total amount of parasite DNA (regardless of number of species present) could be determined. This assay could be used to quantify total *Eimeria* sp. DNA present in a sample but could not be used to separate the various *Eimeria* species in mixed samples. To address this limitation, 7 sets of species-specific PCR primers and probes targeting the mt COI locus were developed to permit multiplexed qPCR assays (with 3 or 4 species per reaction) using HybProbes in a LightCycler® Real-Time PCR Instrument. Theoretically, each species would be amplified by a
different set of primers, and a uniquely-labeled probe could distinguish and quantify each PCR amplicon. Initial attempts using two primer/probe sets to test for *E. acervulina* and *E. tenella* at different concentrations of the two species in a mix were successful. Increasing the number of primer/probe sets to 3 or 4 in a single tube for simultaneous detection became problematic. Lack of specificity in the primer/probe sets or cross talk of emission spectra among detection filters in the instrumentation or both made quantification of more than one species in a mixed DNA sample problematic. Repeated optimization could not overcome these issues. It is concluded that developing a qPCR assay for the closely related *Eimeria* species infecting chicken was near impossible based on the mt COI gene. It is possible that the COI based assay using illumina technology may allow for multiple diagnostic assays for these species.

The phylogeny of apicomplexan parasites remains unresolved (e.g. Barta, 1997; Morrison, 2004). Conflicting morphological and molecular data limited to a few taxa or genetic targets have been largely unreliable for deducing phylogenetic relationships, especially among the closely related species. In most phylogenetic studies, genetic targets were limited to the nuclear 18S rDNA or ribosomal ITS regions (see Tenter et al., 2002). What has been lacking in studies relating to evolutionary development of apicomplexan parasites is the rigorous application of mitochondrial gene or genome sequences in phylogenetic studies with sufficiently large numbers of sampled taxa. In this study, mt COI sequence data (the ‘Barcode region’) and availability of mt genome sequences from divergent apicomplexan parasites was used in conjunction with prior morphological data and nu 18S rDNA sequences to infer phylogenetic relationships among various coccidia. In general, phylogenetic analyses using the COI sequences and concatenated datasets were most useful for deducing evolutionary relationships that were consistent with traditional concepts of relationships among apicomplexan parasites.

Within the family Sarcocystidae the COI and concatenated datasets confirmed the hypothesis that the genus *Hammondia* is paraphyletic (Chapter 11; Ellis et al., 1999; Mugridge et al., 1999; Morrison, 2009). This study suggested that the simplest taxonomic change would be to transfer *H. hammondi* into the genus *Toxoplasma* and reassign a new genotype for the genus *Hammondia* (e.g. *H. heydorni*) so long as such a
change still reflected biological and molecular characteristics for parasites within the Toxoplasmatinae

The coccidian family Eimeriidae is polyphyletic; members of the coccidian family Lankesterellidae clustered within a large clade of eimeriid coccidia currently assigned to the family Eimeriidae. Likewise, many of the eimeriid genera described are paraphyletic. Parasites in the coccidian genera Caryospora, Cyclospora, Eimeria, Isospora and Lankesterella clustered together despite possessing morphologically different oocyst structures; the monophyly of each of the oocyst morphotypes was also not supported suggesting that oocyst morphology is not sufficient as a predictor of evolutionary relatedness (Joyner 1982; Ghimire, 2010). Members of the genus Eimeria Schneider 1875 all possess “tetrasporocystic with dizoic, nonbivalved sporocysts with or without Stieda bodies” (see Upton et al, 1984; Upton, 2000). Cyclospora cayetanensis consistently clustered among Eimeria spp. as reported previously in phylogenetic analyses based on the 18S rRNA gene (e.g. Relman et al., 1996). Although being closely related at both genetic loci, the oocysts of Cyclospora species (2 sporocysts with Stieda bodies, each containing 2 sporozoites; Ortega et al., 1994) differ markedly from those of Eimeria species. Similarly, the morphologically divergent parasites in the genus Caryospora (oocysts with 1 sporocyst with a Stieda body that contains 8 sporozoites) were also found among Eimeria species in phylogenetic reconstructions. Clearly, the genus Eimeria is paraphyletic; the present work reinforces many studies using a variety of genetic loci that have suggested that the genus Eimeria is paraphyletic (Morrison et al., 2004; Martynova-Vankley et al., 2008; Jirků et al., 2009). To resolve this taxonomic challenge, there are two obvious solutions: 1) Lump all coccidia with sporocysts possessing Stieda bodies into a single speciose genus, thus synonymizing the genera Cyclospora, Isospora, Caryospora and Eimeria; or, 2) Reallocate the many described Eimeria species infecting a wide variety of hosts into a number of new genera that could then each be monophyletic. The first solution would still leave a paraphyletic genus Eimeria because of Lankesterella species. Many parasites possessing oocysts without Stieda bodies that were previously placed within with genus Eimeria have been assigned subsequently to other genera (Dyková and Lom, 1981; Upton et al., 1984). Undoubtedly, increasing the number of described genera is more likely to properly represent the
biologically diversity currently residing within the single genus *Eimeria*; however, resolving paraphyly at the genus level may require so many new genera to be erected that such a taxonomic proposal would never be supported by coccidiologists.

Exploration of mitochondrial evolution across the phylum Apicomplexa was possible by generating complete genome sequences from a diverse range of apicomplexan parasites through whole mt genome PCR amplification. During the course of this thesis, complete mt genomes were generated for 6 *Eimeria* spp (5 infecting turkeys and *Eimeria mitis* in chickens), *Cyclospora cayetanensis*, *Caryospora bigenetica*, an unnamed *Isospora* sp. infecting canaries, and *Lankesterella minima* that infects local frogs (see Ogedengbe et al., 2013, 2014, 2015a, 2015b, Ogedengbe and Barta 2015, Ogedengbe et al., unpublished; Chapters 6-10). In addition, partial mt genome sequences (as well as NUMTs and apparent mt minicircles) were obtained for a number of tissue coccidia (i.e. *Toxoplasma gondii*, *Cystoisospora suis* and *C. felis*, Ogedengbe and Barta, unpublished; Chapter 13). These newly generated whole genome sequences were compared with complete mt genomes available from piroplasms, haemosporinid parasites, other *Eimeria* species, and adeleid coccidia (e.g. Hikosaka et al., 2010; Lin et al., 2011; Feagin et al., 2012; Leveille et al., 2015). Collectively, these data were used to elucidate phylogenetic relationships among apicomplexan parasites. Tree topologies based on whole mt genomes (i.e. *Eimeria* spp.), or concatenated protein coding DNA sequences and concatenated amino acid translations for 57 apicomplexan parasites or incomplete mt genome sequences for tissue coccidia supported evolutionary relationships among all the major groups of Apicomplexa consistent with relationships inferred using traditional biological characters, including ultrastructural features. Irrespective of the type of dataset (i.e. concatenated DNA or concatenated amino acid translations of protein coding genes) species considered closely related based on biological characteristics formed well-supported clades and sister groups. In Chapter 13, evolutionary relationships deduced using mt genome sequence datasets resolved relationships readily among closely related species as well as, unexpectedly, much deeper (older) relationships among more distantly related apicomplexan parasites.
Diversity exists in the forms and organization of apicomplexan mt genomes. To date, over 550 typical apicomplexan mt genomes (i.e. ~6kb or greater in length with 3 CDS and rDNA fragments) from 13 genera within four major groups (haemosporinids, piroplasms, eimeriid coccidia, and adeleid coccidia) have been completed; evidence provided by Gjerde (2013) as well as in the present work suggest that only an abbreviated mt genomes exist in the Sarcocystidae. Circular-mapping mt genomes have been reported from haemosporinids (i.e. *Plasmodium*, *Haemoproteus*, *Parahaemoproteus*, *Hepatocystis* and *Leucocytozoon* spp. - see Feagin et al., 1991; Wilson et al., 1992; Feagin, 1992; Preiser et al., 1996; Wilson and Williamson, 1997; Feagin et al., 2000; 2012; Perkins, 2008; Omori et al., 2008); linear monomeric forms with terminal inverted repeats (TIR) in piroplasms (i.e. *Babesia* and *Theileria* spp., Kairo et al., 1994; Omori et al., 2007; Lau, 2009; Hikosaka et al., 2011; 2012). Circular-mapping mt genomes were reported in *Eimeria* spp. infecting chicken (Hikosaka et al., 2010; Lin et al., 2011; Liu et al., 2012). In this study additional complete, circular-mapping mt genomes were sequenced for various eimeriid coccidia (i.e. species in the genera *Eimeria*, *Caryospora*, *Isospora* and *Lankesterella*, Ogedengbe et al., 2013; 2014; Ogedengbe et al., 2015a; Ogedengbe and Barta, 2015; Ogedengbe et al., unpublished data; Chapters 6-10). In the case of the eimeriid coccidium *Cyclospora cayetanensis*, a linear monomeric mt genome with no TIR was reported for this coccidian parasite (see Ogedengbe et al., 2015b) begging the question as to how the single copy mt genome of *C. cayetanensis* manages to replicate itself. Recent completion of the 3′-end of the *C. cayetanensis* mt genome (Cinar et al., 2015) suggests that the 3′-terminus contains repetitive sequences suggestive of telomeric repeats. Of note, an A/T-rich region starting with a ‘‘TATTT’’ motif that is capable of forming a hairpin was found at the 5′-end of the *C. cayetanensis* mt genome just upstream from the SSU/8 rDNA fragment (Ogedengbe et al., 2015b). Apparent homologs of this hairpin region starting with the same TATTT motif were found in all eimeriid and *Plasmodium* mt genomes, although some sequence heterogeneity within the fold region of the loop was present in some of the parasites. We concluded (Ogedengbe et al., 2015b) that this hairpin region identified at the 5′-end of the *C. cayetanensis* mt genome is the likely the ancestral “start” of all
eimeriid mt genomes and suggested that all circular-mapping mt genomes from eimeriid coccidia should use this well-conserved “‘TATTT’” region as the designated 5’-end of each genome to establish a standardised ‘reading frame’ for these eimeriid mt genomes.

In addition to the further reduced mt genomes observed for parasites in the Sarcocystidae (see Gjerde, 2013; Chapters 11-13), even smaller apparent mini-circles were reported for the mt genomes of tissue coccidia analysed in this study (e.g. *Cystoisospora felis* and *C. suis* and *Toxoplasma gondii*, Ogedengbe and Barta, unpublished data). Gjerde (2013) had also reported nuclear mitochondrial (NUMTs) sequences for *T. gondii*, *Hammondia heydorni* and *H. trifittae*, and these apparent NUMTs were detected in the present work as well.

Genome order differs across apicomplexan mt genomes. Typically, transcriptional direction was conserved among species within the widely recognized apicomplexan groups. For example, the protein and rRNA genes were oriented consistently in the same directions in the mt genomes of all eimeriid coccidia examined (Chapters 6-10). Transcriptional directions varied in other groups and were most variable in the piroplasms (Figure 14.1) (Hikosaka et al., 2012). Using the Mauve aligner as a global genome alignment platform, homologous regions across genome sequences were identified as locally collinear blocks (LCB) to illustrate genomes that have undergone rearrangement (see Darling et al., 2004; 2007).

The typical apicomplexan mt genome has conserved gene content with three CDS encoding cytochrome c oxidase subunit I (COI), cytochrome c oxidase subunit
Figure 14.1. The unrooted BI phylogenetic tree representing relationships among major genera of apicomplexan taxa for which mitochondria genome sequences are available was constructed using a concatenation of three protein gene sequences (CytB, COI and COIII). The tree was generated using CDS regions were analyzed using a codon based likelihood model (i.e. GTR+I+G (Nst=6) with Nucmodel=Codon and Code=MetMt). Posterior probabilities of each branch indicate the level of support for the evolutionary relationships illustrated among the various apicomplexan parasites. The simplified mt genome structure for each parasite or group of parasites is found to the right of the phylogenetic tree. The schematic representation illustrates the diverse transcriptional directions found among apicomplexan parasites. Genome organisation is conserved among Plasmodium and relatives (i.e. Leucocytozoon, Haemoproteus, Parahaemoproteus and Hepatocystis); similarly, gene order is conserved in genera within the eimeriid coccidia group (i.e. Eimeria, Cyclospora, Caryospora, Isospora and Lankesterella). Variations in gene order are pronounced in the piroplasms; the four species representative of the group i.e. Babesia microti, B. bigemina, Theileria equi and T. orientalis each exhibited unique directions for the 3 protein coding genes. Terminal inverted repeats (TIR) are illustrated by arrows. The genome order in Hepatozoon is unique to this adeleid parasite; CytB and COIII are inverted in direction and position within this mt genome.
III (COIII) and cytochrome b (CytB) interspersed among highly fragmented large subunit (LSU) and small subunit (SSU) ribosomal RNA genes. Atypically, even more reduced mt genomes appear to exist in tissue coccidia (Sarcocystidae) where only 2 of 3 of the CDS were found in any one parasite. In the case of *Toxoplasma, Neospora* and *Hammondia* species, COI and CytB are encoded whereas in *Cystoisospora* species COI, COIII and a partial CytB are found; in both groups of sarcocystid coccidia, few rDNA fragments were identified. Both gregarine parasites and *Cryptosporidium* species have taken this reduction one step further; they have no apparent mt genome at all (Abrahamsen, 2004; Scicluna, 2006; Mogi and Kita, 2010; Alcock et al., 2012).

The sequences of protein-coding genes in apicomplexan mt genomes were more divergent than the rDNA fragments. The COI gene was the most conserved compared to the COIII and CytB genes in all mt genomes of eimeriid coccidia (Ogedengbe et al., 2013). Genome sizes ranged from 6kb-11kb. Conventional start codons (ATG or ATA or ATT) were used for all COI CDS and the start codon was usually located approximately 30 nucleotides upstream of nucleotides encoding a highly conserved ‘Asn–His–Lys’ amino acid motif. This conserved amino acid motif is associated with the start of the heme–copper oxidase subunit I core region of COI. The putative start codon for COI is frequently found immediately downstream of an A/T-rich region presumed to act as a signal motif (sensu Feagin et al. 1992).

14.2. HIGHLIGHTS:

**Morphological or morphometrics of exogenous stages alone are inadequate to differentiate species (Chapters 11-12)**

Oocyst morphology and morphometrics could not differentiate some closely related coccidia (e.g. *Cystoisospora* species). Morphological identification of *Cystoisospora* spp. was inconclusive due to the comparatively wide range of oocyst measurements reported previously. Among the *Eimeria* species, overlap between distinct *Eimeria* species, even within a single host can complicate identification. For eimeriid and tissue coccidia, well-characterized, single-lines species and samples were used in the present study.
Mitochondrial COI is an effective barcode gene (Chapters 11-12)

COI reference sequences were generated using two sets of degenerate group-specific primers to amplify portions of the mt COI from all eimeriid and all sarcocystid coccidia tested to date. This barcode region, within the mt COI gene, overlaps with the region used in COI-based DNA barcoding initiatives with other taxonomic groups. Collectively, these primer sets amplified DNA from virtually all coccidia. The COI reference sequences and optimized molecular methods developed herein provide a broadly applicable molecular identification method for medically and economically important apicomplexan parasites.

Alternate barcode regions for Cystoisospora species (Chapter 11)

Mitochondrial COI was effective at assigning a particular coccidium specimen to the genus Cystoisospora rather than to another genus of tissue coccidium i.e. Hammondia, Neospora or Toxoplasma. However, the COI locus was unable to differentiate morphologically similar parasites in the genus Cystoisospora. A different DNA barcode locus (perhaps the rapidly evolving nu ribosomal ITS regions or the intergenic regions within their mt genomes) may be required to reliably differentiate Cystoisospora species.

All eimeriid coccidia have conserved mt genome structure and gene content (Chapters 6-10, 13)

Eimeriid mt genomes have three protein-coding genes (COI, COIII and CytB) as found in other apicomplexan parasites. Protein-coding genes were interspersed with numerous regions encoding 15 LSU and 11 SSU rRNA fragments and all coding regions were arranged in the same order and directions in all eimeriid mt genomes examined.

Conserved motifs in the eimeriid coccidian mt genomes indicative of start of CDS (Chapters 6-10, 13)

The putative start codons for CDS in all eimeriid coccidia were frequently found immediately downstream of an A/T-rich region presumed to act as a signal motif. The COI start codon was located approximately 30 nucleotides upstream of nucleotides encoding a highly conserved ‘Asn–His–Lys’ amino acid motif in all coccidian parasites examined in this study.

Most eimeriid coccidia are circular-mapping (Chapters 6-7, 9-10)

Mitochondrial genomes in eimeriid coccidia are found as multiple copies arranged head to tail (linear concatamers) or are circular in structure; thus, genome maps for these parasites are circular-mapping.
A linear monomeric mt genome in *Cyclospora cayetanensis* (Chapter 8)
A linear monomeric mt genome of *C. cayetanensis* was obtained in the present work. Unlike other linear apicomplexan mt genomes, the mt genome of *C. cayetanensis* did not possess terminal inverted repeats. The 5‘-end of the *C. cayetanensis* mt genome has a hypothetical hairpin motif starting with ‘‘TATTT’’ that is conserved in the circular-mapping mt genomes of other coccidia suggesting that this may be the ancestral “start” of all eimeriid mt genomes. It is recommended that this hairpin motif be used to mark the start position of all coccidian mt genomes, regardless if concatenated or circular.

Apicomplexan mt genomes have undergone rearrangement (Chapters 6-10, 13)
Genome order and transcriptional directions were typically conserved within major groups of apicomplexan parasites. Between major groups, changes in transcriptional direction and gene orders were typical. Unusually, transcriptional directions and gene orders were highly varied among piroplasms; sometimes, such variation existed even within a single parasite species e.g. *Babesia microti*.

Typical and atypical mt genomes in Apicomplexa (Chapters 6-11, 13)
Typical mt genomes in Apicomplexa were approximately 6-11kb in size and contained 3 protein-coding genes for COI, COIII and CytB. No such typical apicomplexa mt genomes were detected in tissue coccidia (Sarcocystidae). Instead, circular-mapping, partial mt genomes of 3.5kb were recovered from *Cystoisospora felis* and *C. suis* (containing complete COI and COIII, but only a partial CytB gene). Even shorter 2.7kb mitochondrial sequences (possibly nuclear mitochondrial DNA - NUMTs) were obtained for *Toxoplasma gondii* and *Hammondia heydorni* (containing complete COI and CytB, but with no evidence of a COIII gene). This variability of genome structure within the Sarcocystidae is an area that will require more detailed analysis to determine if a typical mt genome exists in this apicomplexan group.

Mitochondrial genes and genome sequences supported the monophyly of major groups within the phylum Apicomplexa (Chapters 7, 11-13)
Phylogeny based on mitochondrial genes and genome sequences resolved evolutionary relationships at the species level in most cases (exceptionally, not in the case of *Cystoisospora* species). Surprisingly, much deeper (older) relationships among closely or more distantly related apicomplexan parasites could also be inferred when sufficient numbers of taxa were included in the analyses. The evolutionary relationships deduced using mt data were consistent with relationships based on traditional biological characters including ultrastructural features.
REFERENCES


Adl SM, Leander BS, Simpson AGB, Simpson AGB, Archibald JM, Anderson OR, Bass D, Bowser SS, Brugerolle G, Farmer MA, Karpov S, Kolisko M, Lane CE, Lodge DJ, Mann DG, Meisterfeld R, Mendoza L, Moestrup Ø, Mozley-Standridge SE,


Barta JR, Desser SS (1989) Development of Babesia stableri (Dactylosomatidae; Adeleida; Apicomplexa) in its leech vector (Batracobdella picta) and the relationship of the Dactylosomatids to the piroplasms of higher vertebrates. Journal of Protozoology 36:241–253


186


Carreno RA, Martin DS, Barta JR (1999) *Cryptosporidium* is more closely related to the gregarines than to coccidia as shown by phylogenetic analysis of apicomplexan parasites inferred using small-subunit ribosomal RNA gene sequences. Parasitology Research 85:899–904
Carreno RA, Schnitzler BE, Jeffries AC, Tenter AM, Johnson AM, Barta JR (1998) Phylogenetic analysis of coccidia based on 18S rDNA sequence comparison indicates that Isospora is most closely related to Toxoplasma and Neospora. Journal of Eukayotic Microbiology 45:184–188


Clarkson MJ, Gentles MA (1958) Coccidiosis in turkeys. The Veterinary Record 70:211–214


Cushing TL, Schat KA, States SL, Grodio JL, O’Connell PH, Buckles EL (2011) Characterization of the host response in systemic isosporosis (atoxoplasmosis) in a colony of captive American goldfinches (Spinus tristis) and house sparrows (Passer domesticus). Veterinary Pathology 48:985–992

Dahlgren SS, Gjerde B (2010) The red fox (Vulpes vulpes) and the arctic fox (Vulpes lagopus) are definitive hosts of Sarcocystis alces and Sarcocystis hjorti from moose (Alces alces). Parasitology 137:1547–1557


190


Frenkel JK, Dubey JP (1975a) *Hammondia hammondi* gen. nov., sp. nov., from domestic cats, a new coccidium related to *Toxoplasma* and *Sarcocystis*. Zeitschrift für Infectious Diseases 46:3–12


Gjerde B (2013a) Characterisation of full-length mitochondrial copies and partial nuclear copies (numts) of the cytochrome b and cytochrome c oxidase subunit I genes of *Toxoplasma gondii*, *Neospora caninum*, *Hammondia heydorni* and
Hammondia triffittae (Apicomplexa: Sarcocystidae). Parasitology Research 112:1493–1151


Goggin CL, Barker SC (1993) Phylogenetic position of the genus Perkinsus (Protista, Apicomplexa) based on small subunit ribosomal RNA. Molecular and Biochemical Parasitology 60:65–70


caddisflies (Trichoptera). New Zealand Journal of Marine and Freshwater Research 43:1137–1146


Kosa P, Valach M, Tomaska L, Wolfe KH, Nosek J (2006) Complete DNA sequences of the mitochondrial genomes of the pathogenic yeasts *Candida orthopsilosis* and
Candida metapsilosis: Insight into the evolution of linear DNA genomes from mitochondrial telomere mutants. Nucleic Acids Research 34:2472–2481


Levine ND (1985) Veterinary Protozoology. The Iowa State University Press, Ames, Iowa, USA 1–414


Mehlhorn H, Peters W, Haberkorn A (1980) The formation of kinetes and oocyst in Plasmodium gallinaceum (Haemosporidia) and considerations on phylogenetic relationships between Haemosporidia, Piroplasmida and other coccidia Protistologica 16:135–154


Mugridge NB, Morrison DA, Heckeroth AR, Johnson AM, Tenter AM (1999) Phylogenetic analysis based on full length large subunit ribosomal RNA gene sequence comparison reveals that Neospora caninum is more closely related to Hammondia heydorni than to Toxoplasma gondii. International Journal for Parasitology 29:1545–1556
Ogedengbe JD, Hanner RH, Barta JR (2011) DNA barcoding identifies Eimeria species and contributes to the phylogenetics of coccidian parasites (Eimeriorina, Apicomplexa, Alveolata). International Journal for Parasitology 41:843–850


Ogedengbe ME, Barta JR, Ogedengbe JD, Akanbi BO, Ogo IN (2009) Use of a sequence characterized amplified regions (SCARs) multiplex PCR method to identify Eimeria species of chickens from widely distributed geographic samples. Journal of Tropical Veterinarian 27:36–44

Ogedengbe ME, Brash M, Barta JR (2015a) The complete mitochondrial genome sequence of an Isospora sp (Eimeriidae, Eucoccidiorida, Coccidiasina, Apicomplexa) causing systemic coccidiosis in domestic canaries (Serinus canaria Linn). Mitochondria DNA doi 103109/19401736.2015.1018201

Ogedengbe ME, El-Sherry S, Whale J, Barta JR (2014) Complete mitochondrial genome sequences from five Eimeria species (Apicomplexa; Coccidia; Eimeriidae) infecting domestic turkeys Parasite and Vectors 7:335. doi: 101186/1756-3305-7-335


Shirley MW (1975) Enzyme variation in Eimeria species of the chicken Parasitology 71:369–76


Slapeta J (2011) Naming of *Cryptosporidium pestis* is in accordance with the ICZN Code and the name is available for this taxon previously recognised as *C. parvum* "bovine genotype". Veterinary Parasitology 177:1-5


Strüder-Kypke MC, Lynn DH (2010) Comparative analysis of the mitochondrial cytochrome c oxidase subunit 1 (CO1) gene in ciliates (Alveolata, Ciliophora) and evaluation of its suitability as a biodiversity marker. Systematics and Biodiversity 8:131–148


213


Tyzzer EE (1929) The coccidia of the pheasant, turkey, and quail. American Journal of Hygiene 10:1314–1339


Vrba V, Poplstein M, Pakandl M (2011) The discovery of the two types of small subunit ribosomal RNA gene in *Eimeria mitis* contests the existence of *E. mivati* as an independent species. Veterinary Parasitology 183:47–53


APPENDICES

APPENDIX 1: AMINO ACIDS ENCODED BY FOUR MITOCHONDRIAL GENOMES OBTAINED FROM VARIOUS EIMERIA MITIS STRAINS (SEE FIGURES 6.2 AND 6.3 FOR AN EXPLANATION). AMINO ACIDS THAT DIFFERED IN FREQUENCY AMONG THE FOUR GENOMES ARE IN BOLD TEXT.

<table>
<thead>
<tr>
<th>Amino Acids</th>
<th>E. mitis 24</th>
<th>E. mitis 20</th>
<th>E. mitis 18</th>
<th>E. mitis (China)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KC409031</td>
<td>KC409030</td>
<td>KC409029</td>
<td>JN864949</td>
</tr>
<tr>
<td></td>
<td>Freq</td>
<td>%</td>
<td>Freq</td>
<td>%</td>
</tr>
<tr>
<td>Alanine</td>
<td>59</td>
<td>54%</td>
<td>58</td>
<td>53%</td>
</tr>
<tr>
<td>Cysteine</td>
<td>15</td>
<td>14%</td>
<td>15</td>
<td>14%</td>
</tr>
<tr>
<td>Aspartic Acid</td>
<td>18</td>
<td>16%</td>
<td>18</td>
<td>16%</td>
</tr>
<tr>
<td>Glutamic Acid</td>
<td>21</td>
<td>19%</td>
<td>22</td>
<td>20%</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>98</td>
<td>90%</td>
<td>98</td>
<td>90%</td>
</tr>
<tr>
<td>Glycine</td>
<td>79</td>
<td>72%</td>
<td>77</td>
<td>70%</td>
</tr>
<tr>
<td>Histidine</td>
<td>25</td>
<td>23%</td>
<td>25</td>
<td>23%</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>103</td>
<td>94%</td>
<td>101</td>
<td>92%</td>
</tr>
<tr>
<td>Lysine</td>
<td>13</td>
<td>12%</td>
<td>13</td>
<td>12%</td>
</tr>
<tr>
<td>Leucine</td>
<td>168</td>
<td>154%</td>
<td>168</td>
<td>154%</td>
</tr>
<tr>
<td>Methionine</td>
<td>25</td>
<td>23%</td>
<td>26</td>
<td>24%</td>
</tr>
<tr>
<td>Asparagine</td>
<td>36</td>
<td>33%</td>
<td>36</td>
<td>33%</td>
</tr>
<tr>
<td>Proline</td>
<td>39</td>
<td>36%</td>
<td>39</td>
<td>36%</td>
</tr>
<tr>
<td>Glutamine</td>
<td>21</td>
<td>19%</td>
<td>21</td>
<td>19%</td>
</tr>
<tr>
<td>Arginine</td>
<td>25</td>
<td>23%</td>
<td>25</td>
<td>23%</td>
</tr>
<tr>
<td>Serine</td>
<td>104</td>
<td>95%</td>
<td>105</td>
<td>96%</td>
</tr>
<tr>
<td>Threonine</td>
<td>81</td>
<td>74%</td>
<td>82</td>
<td>75%</td>
</tr>
<tr>
<td>Valine</td>
<td>79</td>
<td>72%</td>
<td>80</td>
<td>73%</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>23</td>
<td>21%</td>
<td>23</td>
<td>21%</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>58</td>
<td>53%</td>
<td>58</td>
<td>53%</td>
</tr>
<tr>
<td>STOP</td>
<td>*</td>
<td>03%</td>
<td>3</td>
<td>03%</td>
</tr>
</tbody>
</table>
Supplementary Figure S1

Supplementary Figure S7.1. - Bayesian inference (BI) and maximum likelihood (ML) phylogenetic reconstructions based on aligned concatenated CDS of COI, COIII and CytB from the mitochondrial genomes of 16 Eimeria species. The analyses included 5 species infecting turkeys and 7 species infecting chickens and used Eimeria magna (a parasite of rabbits) as the functional outgroup to root the tree. Node support is indicated for BI (posterior probability, first number) and for ML (% bootstrap, second number) for all nodes with greater than 0.5 posterior probabilities. Neither the Eimeria species infecting chickens nor the Eimeria species infecting turkeys formed monophyletic groups. Both the BI and ML analyses supported monophyly of the 5 Eimeria species of chickens that do not usually invade the cecal pouches but branching order among these parasites was poorly resolved in both. The same tree topology was obtained based on aligned near-complete mitochondrial genome sequences (see Figure 7.2).
**Supplementary Figure S2**

Maximum parsimony phylogenetic reconstruction based on aligned concatenated CDS of COI, COIII and CytB from the mitochondrial genomes of 16 *Eimeria* species. The analyses included 5 species infecting turkeys and 7 species infecting chickens and used *Eimeria magna* (a parasite of rabbits) as the functional outgroup to root the tree. Percentage bootstrap support (500 replicates) is indicated at each node with at least 50% bootstrap support. The MP tree differed from the BI/ML tree only in the placement of *E. meleagrimitis* basal to a collection of lower intestinal tract parasites of chickens and turkeys. The MP analysis supported monophyly of the 5 *Eimeria* species of chickens that do not usually invade the cecal pouches. The same tree topology was obtained based on aligned near-complete mitochondrial genome sequences (see Figure 7.3).
### APPENDIX 3: LIST OF SEQUENCES, STRAINS AND ACCESSION NUMBERS USED IN CHAP1ER 12

<table>
<thead>
<tr>
<th>SPECIES (SPECIES (</th>
<th>18S rDNA GENBANK NUMBERS</th>
<th>SPECIES (SPECIES (</th>
<th>COI GENBANK NUMBERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isospora sp ex Serinus canaria</td>
<td>KT184357</td>
<td>Isospora sp ex Serinus canaria</td>
<td>KP658103</td>
</tr>
<tr>
<td>Atoxoplasma MS-2003 50%</td>
<td>KY331571, KY331569</td>
<td>Atoxoplasma MS-2003 50%</td>
<td>KP658103</td>
</tr>
<tr>
<td>Caryospora bigenicuta 50%</td>
<td>AF060975, AF060976, KT184331, KT184332</td>
<td>Caryospora bigenicuta 50%</td>
<td>KF859856, KP658102</td>
</tr>
<tr>
<td>Cyclospora caytanensis 50%</td>
<td>AF111183, U40261</td>
<td>Cyclospora caytanensis</td>
<td>KP658101</td>
</tr>
<tr>
<td>Cystoisospora felis 50%</td>
<td>LC6471, KC807826, KT184364</td>
<td>Cystoisospora felis</td>
<td>JN227478</td>
</tr>
<tr>
<td>Eimeria acalcula 50%</td>
<td>KT184334, AF33850</td>
<td>Eimeria acalcula</td>
<td>KT184373</td>
</tr>
<tr>
<td>Eimeria bovis 90%</td>
<td>AB769574, AB769575, AB769576, AB769577, AB769578, AB769579, AB769580, AB769581, AB769582, AB769583, AB769584, AB769585, AB769586, AB769587, AB769589, KT184336</td>
<td>Eimeria bovis</td>
<td>KT184372</td>
</tr>
<tr>
<td>Eimeria acalcula 50%</td>
<td>U67115, DQ31687, FJ236372, DQ583651, EF210324, EF175928, EF210323, KT184333</td>
<td>Eimeria acalcula 50%</td>
<td>HQ702479, HM771674, HM771673, FJ216644, FJ216648, FJ216642, FJ216640, FJ185855</td>
</tr>
<tr>
<td>Eimeria adenoeides 90%</td>
<td>KC305170, KC305171, KC305172, KC305173, KC305174, KC305175, KC305176, KC305177, KC305178, KC305179, KC305180, KC305181, KC305182, KC305183, KC305184, KC305185, KC305169, FR745914, HG793240</td>
<td>Eimeria adenoeides 50%</td>
<td>KC46358, KC46359, KC46360, KC46361, FR464202, HM117017, HM117018, HG793047</td>
</tr>
<tr>
<td>Eimeria alabamensis 90%</td>
<td>AB769549, AB769550, AB769551, AB769552, AB769553, AB769554, AB769555, AB769556, AB769547, AB769548, AF291242, KT184335</td>
<td>Eimeria alabamensis</td>
<td>KT184376</td>
</tr>
<tr>
<td>Eimeria alorani</td>
<td>JQ993639</td>
<td>Eimeria alorani</td>
<td>JQ993701</td>
</tr>
<tr>
<td>Eimeria bruneti 50%</td>
<td>U67116, KT184337</td>
<td>Eimeria bruneti 50%</td>
<td>HM771675, HQ702480</td>
</tr>
<tr>
<td>Eimeria hardi</td>
<td>JQ993666</td>
<td>Eimeria hardi</td>
<td>JQ993709</td>
</tr>
<tr>
<td>Eimeria cahirinsensis 50%</td>
<td>JQ993647, JQ993648</td>
<td>Eimeria cahirinsensis 50%</td>
<td>JQ993687, JQ993686</td>
</tr>
<tr>
<td>Eimeria caviae</td>
<td>JQ993649</td>
<td>Eimeria caviae</td>
<td>JQ993689</td>
</tr>
<tr>
<td>Eimeria cf. tenggilingi</td>
<td>JX464223</td>
<td>Eimeria cf. tenggilingi</td>
<td>JX464222</td>
</tr>
<tr>
<td>Eimeria coecocilica 50%</td>
<td>HQ173828, EF694015</td>
<td>Eimeria coecocilica</td>
<td>JQ993690</td>
</tr>
<tr>
<td>Eimeria dispersa Briston 50%</td>
<td>HQ173841, KT184338</td>
<td>Eimeria dispersa 50%</td>
<td>HG793048, KJ608416</td>
</tr>
<tr>
<td>Eimeria exiguus 50%</td>
<td>HQ173829, EF694007</td>
<td>Eimeria exiguus</td>
<td>JQ993691</td>
</tr>
<tr>
<td>Eimeria falciformis 50%</td>
<td>AF080614, KT184339, KT184340</td>
<td>Eimeria falciformis</td>
<td>HM771682</td>
</tr>
<tr>
<td>Eimeria flavescent 50%</td>
<td>HQ173830, EF694011</td>
<td>Eimeria flavescent</td>
<td>JQ993692</td>
</tr>
<tr>
<td>Eimeria gallopavonis 50%</td>
<td>HM117011, HG793042, KT184341, KT184342, KT184343, KT184344</td>
<td>Eimeria gallopavonis 50%</td>
<td>HM117018, HG793051, J608413, J608417</td>
</tr>
<tr>
<td>Eimeria innocua</td>
<td>HG793045</td>
<td>Eimeria innocua</td>
<td>HG793049</td>
</tr>
<tr>
<td>Eimeria intestinal 50%</td>
<td>HQ173831, EF694012</td>
<td>Eimeria intestinal</td>
<td>JQ993693</td>
</tr>
<tr>
<td>Eimeria irussidua 50%</td>
<td>HQ173827, EF694009</td>
<td>Eimeria irussidua</td>
<td>JQ993694</td>
</tr>
<tr>
<td>Eimeria macrospodis 50%</td>
<td>JQ392575, JQ392576, JQ392577</td>
<td>Eimeria macrospodis 50%</td>
<td>JQ392578, JQ392579, JQ392580</td>
</tr>
<tr>
<td>Eimeria magna 50%</td>
<td>HQ173833, EF694016</td>
<td>Eimeria magna</td>
<td>JQ993695</td>
</tr>
<tr>
<td>Eimeria maxima Type B 50%</td>
<td>FJ236329, FJ236330, FJ236332, FJ236333, DQ316816, DQ583650, EU025110, DQ583648, KT184345</td>
<td>Eimeria maxima 90%</td>
<td>HQ702481, FJ236387, FJ236389, FJ236390, FJ236393, FJ236402, FJ236406, FJ236412, FJ236417, FJ236418, FJ236435, FJ236437, FJ236440, FJ236448, FJ236459, HM771684</td>
</tr>
<tr>
<td>Eimeria meleagridis 50%</td>
<td>FR745913, KT184347, KT184348</td>
<td>Eimeria meleagridis 50%</td>
<td>FR846201, HG793046, J608418</td>
</tr>
<tr>
<td>Eimeria mycteria 50%</td>
<td>JF301418</td>
<td>Eimeria mycteria 50%</td>
<td>JQ993706</td>
</tr>
<tr>
<td>Eimeria nafako</td>
<td>JQ993665</td>
<td>Eimeria nafako</td>
<td>JQ993708</td>
</tr>
<tr>
<td>Eimeria meleagrimitis Type B 50%</td>
<td>KC305187, KC305194, KC305196, KC305197, HG793043, HM117012</td>
<td>Eimeria meleagrimitis 50%</td>
<td>KC46351, KC46352, KC46353, KC46354, J608414, HG793050,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPECIES ( % concensus)</td>
<td>18S rDNA (GENBANK NUMBERS)</td>
<td>SPECIES ( % concensus)</td>
<td>COI (GENBANK NUMBERS)</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------------</td>
<td>------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Eimeria meleagrimitis Type A 90%</td>
<td>KC305190, KC305191, KC305192, KC305193, KC305195, KC305198, KC305199, KC305186, KC305189, HG793044, HG793043, HM117013, HM117014</td>
<td>Eimeria meleagrimitis 50%</td>
<td>KC346351, KC346352, KC346353, KC346354, KJ608414, HG793050,</td>
</tr>
<tr>
<td>Eimeria meleagrimitis Type A 90%</td>
<td>U76748, FJ236377, R775305, FJ236378, FR775307, FR775306, FJ236374, FJ236373, FJ236375, FJ236376</td>
<td>Eimeria meleagrimitis 50%</td>
<td>JN864949, FJ236433, FJ236434, FR796699, HM771681, EY174185, KC409030, KP501573</td>
</tr>
<tr>
<td>Eimeria meleagrimitis Type B 50%</td>
<td>FJ236379, FR775304, FR775303, FR775302, U40262, U67118</td>
<td>Eimeria meleagrimitis 50%</td>
<td>JN864949, FJ236433, FJ236434, FR796699, HM771681, EY174185, KC409030, KP501573</td>
</tr>
<tr>
<td>Eimeria necatrix 50%</td>
<td>DQ136185, U67119, EU044765, EU044766, EU044767, KT184349</td>
<td>Eimeria necatrix 50%</td>
<td>HQ702482, HM771680, EU025108</td>
</tr>
<tr>
<td>Eimeria papionina</td>
<td>JN596589</td>
<td>Eimeria papionina</td>
<td>JN596590</td>
</tr>
<tr>
<td>Eimeria papillata</td>
<td>KT184350</td>
<td>Eimeria papillata</td>
<td>KT184377</td>
</tr>
<tr>
<td>Eimeria piriformis 50%</td>
<td>EF694014, HQ173836</td>
<td>Eimeria piriformis</td>
<td>JQ993608</td>
</tr>
<tr>
<td>Eimeria praecox Type B 50%</td>
<td>FJ236367, FJ236368, FJ236369, FJ236370, FJ236371, GQ421692, KT184351</td>
<td>Eimeria praecox 50%</td>
<td>JQ659301, HQ702483, KT184378</td>
</tr>
<tr>
<td>Eimeria sp. ex. Alectoris graeca</td>
<td>HM070377</td>
<td>Eimeria sp. ex. Alectoris graeca</td>
<td>HM117020</td>
</tr>
<tr>
<td>Eimeria sp. ex. Apodemus agrarius 21617</td>
<td>JQ993658</td>
<td>Eimeria sp. ex. Apodemus agrarius</td>
<td>JQ993700</td>
</tr>
<tr>
<td>Eimeria sp. ex. Apodemus agrarius 21668</td>
<td>JQ993660</td>
<td>Eimeria sp. ex. Apodemus agrarius</td>
<td>JQ993702</td>
</tr>
<tr>
<td>Eimeria sp. ex. Apodemus sylvaticus 08-50</td>
<td>JQ993661</td>
<td>Eimeria sp. ex. Apodemus sylvaticus 08-50</td>
<td>JQ993706</td>
</tr>
<tr>
<td>Eimeria sp. ex. Apodemus sylvaticus 08-53</td>
<td>JQ993662</td>
<td>Eimeria sp. ex. Apodemus sylvaticus 08-53</td>
<td>JQ993707</td>
</tr>
<tr>
<td>Eimeria sp. ex. Phataginus tricuspis</td>
<td>JQ993651</td>
<td>Eimeria sp. ex. Phataginus tricuspis</td>
<td>JQ993697</td>
</tr>
<tr>
<td>Eimeria sp. ex. Phasianus colchicus 50%</td>
<td>HM117008, HM117007</td>
<td>Eimeria sp. ex. Phasianus colchicus</td>
<td>HM117019</td>
</tr>
<tr>
<td>Eimeria tamiasciuri</td>
<td>KT184353</td>
<td>Eimeria tamiasciuri</td>
<td>KT184375</td>
</tr>
<tr>
<td>Eimeria trichosuri 50%</td>
<td>FJ829320, FJ829321, FJ829322, FJ829323</td>
<td>Eimeria trichosuri</td>
<td>JN192136</td>
</tr>
<tr>
<td>Eimeria vejdovskyi 50%</td>
<td>HQ173838, EP694010</td>
<td>Eimeria vejdovskyi</td>
<td>JQ993699</td>
</tr>
<tr>
<td>Eimeria zuernii 90%</td>
<td>AY876932, AB769655, AB769656, AB769657, AB769658, AB769659, AB769660, AB769661, AB769662, AB769663, AB769664, AB769665, KT184356</td>
<td>Eimeria zuernii 90%</td>
<td>HM771687</td>
</tr>
<tr>
<td>Eimeria zuernii 50%</td>
<td>AY876932, AB769655, AB769656, AB769657, AB769658, AB769659, AB769660, AB769661, AB769662, AB769663, AB769664, AB769665, KT184356</td>
<td>Eimeria zuernii 50%</td>
<td>HM771687</td>
</tr>
<tr>
<td>Isospora gryphoni 50%</td>
<td>AF080613, KF854254</td>
<td>Isospora gryphoni</td>
<td>KC346355</td>
</tr>
<tr>
<td>SPECIES (GENBANK NUMBERS)</td>
<td>SPECIES (GENBANK NUMBERS)</td>
<td>COI (GENBANK NUMBERS)</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------------------------</td>
<td>-----------------------</td>
<td></td>
</tr>
<tr>
<td>Isospora sp. ex Talpa europaea - JQ993630</td>
<td>Isospora sp. ex Talpa europaea - JQ993713</td>
<td>Isospora sp. ex Talpa europaea - JQ993703</td>
<td></td>
</tr>
<tr>
<td>Isospora sp. ex Talpa europaea - JQ993669</td>
<td>Isospora sp. ex Talpa europaea - JQ993713</td>
<td>Isospora sp. ex Talpa europaea - JQ993703</td>
<td></td>
</tr>
<tr>
<td>Isospora superbusi-Genotype B - KF648870</td>
<td>Isospora superbusi - KF648869</td>
<td>Isospora superbusi - KF648869</td>
<td></td>
</tr>
<tr>
<td>Isospora greineri-Genotype A - KF648870</td>
<td>Isospora greineri - KF648868</td>
<td>Isospora greineri - KF648868</td>
<td></td>
</tr>
<tr>
<td>Lankesterella minima - KT184358, AF080611 50%</td>
<td>Lankesterella minima - KT184381, KT184382</td>
<td>Lankesterella minima - KT184381, KT184382</td>
<td></td>
</tr>
<tr>
<td>Neospora caninum 50% - AJ271354, U16159, U17346, GQ892006, U03069, L24380</td>
<td>Neospora caninum 50% - JX473252, HM771688</td>
<td>Neospora caninum 50% - JX473252, HM771688</td>
<td></td>
</tr>
<tr>
<td>Toxoplasma gondii 50% - U17348, L37415, EF72967, L24381, M97705</td>
<td>Toxoplasma gondii 50% - KF854274, JX473253, HM771670, HM771689</td>
<td>Toxoplasma gondii 50% - KF854274, JX473253, HM771670, HM771689</td>
<td></td>
</tr>
</tbody>
</table>
### APPENDIX 4: LIST OF ALL NEW SEQUENCES SUBMITTED TO GENBANK (NOT USED IN THESIS CHAPTERS)

<table>
<thead>
<tr>
<th>NAME</th>
<th>STRAIN</th>
<th>SOURCE</th>
<th>GENE</th>
<th>ACCESSION NUMBERS</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Isospora</em> sp</td>
<td>JRB-2013</td>
<td>Feacal Sample</td>
<td>Plastid LSU</td>
<td>KF815278</td>
</tr>
<tr>
<td><em>Isospora</em> sp</td>
<td>JRB-2013</td>
<td>Feacal Sample</td>
<td>Plastid SSU</td>
<td>KF815279</td>
</tr>
<tr>
<td>Caryospora cf bigenetica</td>
<td>MTZ_07</td>
<td>Feacal Sample</td>
<td>Plastid LSU</td>
<td>KF859855</td>
</tr>
<tr>
<td>Caryospora cf bigenetica</td>
<td>MTZ_07</td>
<td>Feacal Sample</td>
<td>Plastid SSU</td>
<td>KF859857</td>
</tr>
<tr>
<td><em>Eimeria</em> sp. ex. <em>Oryctolagus cuniculus</em></td>
<td>MEO2015</td>
<td>Feacal Sample</td>
<td>Mt COI</td>
<td>KT184374</td>
</tr>
<tr>
<td><em>Eimeria</em> sp. ex. <em>Branta canadensis</em></td>
<td>MEO2015a</td>
<td>Feacal Sample</td>
<td>Mt COI</td>
<td>KT184379</td>
</tr>
<tr>
<td><em>Eimeria</em> sp. ex. <em>Branta canadensis</em></td>
<td>MEO2015b</td>
<td>Feacal Sample</td>
<td>Mt COI</td>
<td>KT184380</td>
</tr>
</tbody>
</table>
**APPENDIX 5: LIST OF ALL NEW SEQUENCES USED IN THE CHAPTERS OF THE THESIS ONLY, SUBMITTED TO GenBank**

<table>
<thead>
<tr>
<th>Name</th>
<th>Strain</th>
<th>Source</th>
<th>Gene</th>
<th>Accession Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cystoisospora suis</td>
<td>MEO2010a</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KF854251</td>
</tr>
<tr>
<td>Cystoisospora suis</td>
<td>MEO2010b</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KF854252</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>MEO2015</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184364</td>
</tr>
<tr>
<td>Cystoisospora canis</td>
<td>MEO2015a</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184368</td>
</tr>
<tr>
<td>Cystoisospora canis</td>
<td>MEO2015b</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184363</td>
</tr>
<tr>
<td>Cystoisospora canis</td>
<td>MEO2015c</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184361</td>
</tr>
<tr>
<td>Cystoisospora canis</td>
<td>MEO2015d</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184360</td>
</tr>
<tr>
<td>Cystoisospora canis</td>
<td>MEO2015e</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184362</td>
</tr>
<tr>
<td>Cystoisospora ohioensis</td>
<td>MEO2015a</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184359</td>
</tr>
<tr>
<td>Cystoisospora ohioensis</td>
<td>MEO2015b</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184365</td>
</tr>
<tr>
<td>Cystoisospora ohioensis</td>
<td>MEO2015c</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184366</td>
</tr>
<tr>
<td>Cystoisospora rivolta</td>
<td>MEO2015</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184367</td>
</tr>
<tr>
<td>Hammondia hammondi</td>
<td>H-H20</td>
<td>DNA Template</td>
<td>18S rDNA</td>
<td>KT184369</td>
</tr>
<tr>
<td>Hammondia heydorni</td>
<td>MEO2015</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184370</td>
</tr>
<tr>
<td>Sarcocystis neurona</td>
<td>RTTW</td>
<td>Tissue Specimen</td>
<td>18S rDNA</td>
<td>KT184371</td>
</tr>
<tr>
<td>Caryospora cf. bigenetica</td>
<td>MEO2015a</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184331</td>
</tr>
<tr>
<td>Caryospora cf. bigenetica</td>
<td>MEO2015b</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184332</td>
</tr>
<tr>
<td>Eimeria acervulina</td>
<td>NC3</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184333</td>
</tr>
<tr>
<td>Eimeria ahsata</td>
<td>MEO2015</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184334</td>
</tr>
<tr>
<td>Eimeria alabamensis</td>
<td>MEO2015</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184335</td>
</tr>
<tr>
<td>Eimeria bovis</td>
<td>MEO2015</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184336</td>
</tr>
<tr>
<td>Eimeria brunetti</td>
<td>Guelph 80</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184337</td>
</tr>
<tr>
<td>Eimeria dispersa</td>
<td>Briston</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184338</td>
</tr>
<tr>
<td>Eimeria falciformis</td>
<td>Chob2A</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184339</td>
</tr>
<tr>
<td>Eimeria falciformis</td>
<td>Chob2</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184340</td>
</tr>
<tr>
<td>Eimeria gallopavonis</td>
<td>USKS06-01 (Type A)</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184341</td>
</tr>
<tr>
<td>Eimeria gallopavonis</td>
<td>USKS06-01 (Type B)</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184342</td>
</tr>
<tr>
<td>Eimeria gallopavonis</td>
<td>Weybridge (Type A)</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184343</td>
</tr>
<tr>
<td>Eimeria gallopavonis</td>
<td>Weybridge (Type B)</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184344</td>
</tr>
<tr>
<td>Eimeria maxima</td>
<td>GS-74</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184345</td>
</tr>
<tr>
<td>Eimeria maxima</td>
<td>M6</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184346</td>
</tr>
<tr>
<td>Eimeria meleagris</td>
<td>USAR79-01 (Type A)</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184347</td>
</tr>
<tr>
<td>Eimeria meleagris</td>
<td>USAR79-01 (Type B)</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184348</td>
</tr>
<tr>
<td>Eimeria necatrix</td>
<td>Guelph_84</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184349</td>
</tr>
<tr>
<td>Eimeria papillata</td>
<td>Chobo3</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184350</td>
</tr>
<tr>
<td>Eimeria praecox</td>
<td>MEO2015a</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184351</td>
</tr>
<tr>
<td>Eimeria praecox</td>
<td>MEO2015b</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184352</td>
</tr>
<tr>
<td>Eimeria tamiaisciri</td>
<td>MEO2015</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184353</td>
</tr>
<tr>
<td>Eimeria tenella</td>
<td>USDA_80</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184354</td>
</tr>
<tr>
<td>Eimeria vermiformis</td>
<td>Chob1</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184355</td>
</tr>
<tr>
<td>Eimeria zuernii</td>
<td>MEO2015</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184356</td>
</tr>
<tr>
<td>Isospora sp.</td>
<td>MEO2015-xd24</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KT184357</td>
</tr>
<tr>
<td>Isospora gryphoni</td>
<td>UWO2010</td>
<td>Feacal Sample</td>
<td>18S rDNA</td>
<td>KF854254</td>
</tr>
<tr>
<td>Lankesterella minima</td>
<td>MEO2015</td>
<td>Blood Sample</td>
<td>18S rDNA</td>
<td>KT184358</td>
</tr>
<tr>
<td>Sarcocystis neurona</td>
<td>TWC2013</td>
<td>Tissue Specimen</td>
<td>18S rDNA</td>
<td>KF854255</td>
</tr>
<tr>
<td>NAME</td>
<td>STRAIN</td>
<td>SOURCE</td>
<td>GENE</td>
<td>ACCESSION NUMBERS</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>--------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Hammondia hammondi</td>
<td>H H-20</td>
<td>DNA Template</td>
<td>18S rDNA</td>
<td>KF854253</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>MEO2010b</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854257</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>MEO2010a</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854256</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>MEO2010c</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854258</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>MEO2010d</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854259</td>
</tr>
<tr>
<td>Cystoisospora felis</td>
<td>MEO2015-7</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KT184386</td>
</tr>
<tr>
<td>Cystoisospora ohioensis</td>
<td>MEO2010a</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854260</td>
</tr>
<tr>
<td>Cystoisospora ohioensis</td>
<td>MEO2010b</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854261</td>
</tr>
<tr>
<td>Cystoisospora ohioensis</td>
<td>MEO2010c</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854265</td>
</tr>
<tr>
<td>Cystoisospora ohioensis</td>
<td>MEO2010d</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KT184384</td>
</tr>
<tr>
<td>Cystoisospora suis</td>
<td>MEO2010a</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854262</td>
</tr>
<tr>
<td>Cystoisospora suis</td>
<td>MEO2010b</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854263</td>
</tr>
<tr>
<td>Cystoisospora suis</td>
<td>MEO2013a</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854264</td>
</tr>
<tr>
<td>Cystoisospora suis</td>
<td>MEO2013b</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854265</td>
</tr>
<tr>
<td>Cystoisospora canis</td>
<td>ME2015-Milly</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KT184385</td>
</tr>
<tr>
<td>Cystoisospora rivolta</td>
<td>MEO2015-IDX443</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KT184383</td>
</tr>
<tr>
<td>Hammondia hammondi</td>
<td>H H-20</td>
<td>DNA Template</td>
<td>mtCOI</td>
<td>KT184387</td>
</tr>
<tr>
<td>Hammondia hammondi</td>
<td>H H-20</td>
<td>DNA Template</td>
<td>mtCOI</td>
<td>KF854267</td>
</tr>
<tr>
<td>Hammondia heydorni</td>
<td>MEO2010a</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854268</td>
</tr>
<tr>
<td>Hammondia heydorni</td>
<td>MEO2010b</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854269</td>
</tr>
<tr>
<td>Hammondia heydorni</td>
<td>MEO2010e</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854270</td>
</tr>
<tr>
<td>Hammondia heydorni</td>
<td>MEO2015</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KT184388</td>
</tr>
<tr>
<td>Hammondia sp</td>
<td>MEO2010c</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854271</td>
</tr>
<tr>
<td>Sarcocystis neurona</td>
<td>BR2012</td>
<td>Tissue Specimen</td>
<td>mtCOI</td>
<td>KF854272</td>
</tr>
<tr>
<td>Sarcocystis rileyi</td>
<td>MEO2015-08131</td>
<td>Tissue Specimen</td>
<td>mtCOI</td>
<td>KT184389</td>
</tr>
<tr>
<td>Toxoplasma gondii</td>
<td>CTG</td>
<td>DNA Template</td>
<td>mtCOI</td>
<td>KF854273</td>
</tr>
<tr>
<td>Toxoplasma gondii</td>
<td>GT1</td>
<td>DNA Template</td>
<td>mtCOI</td>
<td>KF854274</td>
</tr>
<tr>
<td>Toxoplasma gondii</td>
<td>MAS</td>
<td>DNA Template</td>
<td>mtCOI</td>
<td>KF854275</td>
</tr>
<tr>
<td>Toxoplasma gondii</td>
<td>PTG</td>
<td>DNA Template</td>
<td>mtCOI</td>
<td>KF854276</td>
</tr>
<tr>
<td>Toxoplasma gondii</td>
<td>TgCatBr5</td>
<td>DNA Template</td>
<td>mtCOI</td>
<td>KF854277</td>
</tr>
<tr>
<td>Toxoplasma gondii</td>
<td>TgCatBr64</td>
<td>DNA Template</td>
<td>mtCOI</td>
<td>KF854278</td>
</tr>
<tr>
<td>Toxoplasma gondii</td>
<td>TgCgCal</td>
<td>DNA Template</td>
<td>mtCOI</td>
<td>KF854279</td>
</tr>
<tr>
<td>Toxoplasma gondii</td>
<td>TgToucan</td>
<td>DNA Template</td>
<td>mtCOI</td>
<td>KF854280</td>
</tr>
<tr>
<td>Eimeria ahsata</td>
<td>ME2015</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KT184373</td>
</tr>
<tr>
<td>Eimeria alabamensis</td>
<td>MEO2015</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KT184376</td>
</tr>
<tr>
<td>Eimeria bovis</td>
<td>MEO2015</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KT184372</td>
</tr>
<tr>
<td>Eimeria praecox</td>
<td>MEO2015</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KT184378</td>
</tr>
<tr>
<td>Eimeria papillata</td>
<td>Chob 3</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KT184377</td>
</tr>
<tr>
<td>Eimeria tamiasciuri</td>
<td>MEO2015</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KT184375</td>
</tr>
<tr>
<td>Eimeria sp</td>
<td>MEO2010</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF854266</td>
</tr>
<tr>
<td>Caryospora sp</td>
<td>MTZ_07</td>
<td>Feecal Sample</td>
<td>mtCOI</td>
<td>KF859856</td>
</tr>
<tr>
<td>Lankesterella minima</td>
<td>MEO2015a</td>
<td>Blood Sample</td>
<td>mtCOI</td>
<td>KT184381</td>
</tr>
<tr>
<td>Lankesterella minima</td>
<td>MEO2015b</td>
<td>Blood Sample</td>
<td>mtCOI</td>
<td>KT184382</td>
</tr>
<tr>
<td>Eimeria mitis</td>
<td>USDA 50</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>KF5011573</td>
</tr>
<tr>
<td>Eimeria mitis</td>
<td>USDA 50</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>KC409029</td>
</tr>
<tr>
<td>Eimeria mitis</td>
<td>USDA 50</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>KC409030</td>
</tr>
<tr>
<td>Eimeria mitis</td>
<td>USDA 50</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>KC409031</td>
</tr>
<tr>
<td>NAME</td>
<td>STRAIN</td>
<td>SOURCE</td>
<td>GENE</td>
<td>ACCESSION NUMBERS</td>
</tr>
<tr>
<td>-------------------------</td>
<td>----------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>-------------------</td>
</tr>
<tr>
<td><em>Eimeria dispersa</em></td>
<td>Briston</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>KJ608416</td>
</tr>
<tr>
<td><em>Eimeria meleagrimitis</em></td>
<td>USMN08-01</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>KJ608414</td>
</tr>
<tr>
<td><em>Eimeria adenoeides</em></td>
<td>Guelph</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>KJ608415</td>
</tr>
<tr>
<td><em>Eimeria meleagridis</em></td>
<td>USA R97-01</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>KJ608418</td>
</tr>
<tr>
<td><em>Eimeria gallopavonis</em></td>
<td>Weybridge</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>KJ608413</td>
</tr>
<tr>
<td><em>Eimeria gallopavonis</em></td>
<td>USKS06-01</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>KJ608417</td>
</tr>
<tr>
<td><em>Caryospora bigenetica</em></td>
<td>MTZ_07-12108</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>KP658102</td>
</tr>
<tr>
<td><em>Isospora sp ex Serinus</em></td>
<td>dx23/dx24</td>
<td>DNA Template</td>
<td>mt genome</td>
<td>KP658103</td>
</tr>
<tr>
<td><em>Cyclospora cayetanensis</em></td>
<td>Cyclo_CDC_2013</td>
<td>DNA Template</td>
<td>mt genome</td>
<td>KP658101</td>
</tr>
<tr>
<td><em>Lankesterella minima</em></td>
<td>MEO2015a</td>
<td>Blood Sample</td>
<td>mt genome</td>
<td>KT369005</td>
</tr>
<tr>
<td><em>Lankesterella minima</em></td>
<td>MEO2015b</td>
<td>Blood Sample</td>
<td>mt genome</td>
<td>KT369006</td>
</tr>
<tr>
<td><em>Cystoisospora felis</em></td>
<td>MEO2015</td>
<td>Feecal Sample</td>
<td>mini-genome</td>
<td>KT363922</td>
</tr>
<tr>
<td><em>Cystoisospora suis</em></td>
<td>MEO2015</td>
<td>Feecal Sample</td>
<td>mini-genome</td>
<td>KT363923</td>
</tr>
<tr>
<td><em>Toxoplasma gondii</em></td>
<td>GT1</td>
<td>DNA Template</td>
<td>mini-genome</td>
<td>KT363924</td>
</tr>
<tr>
<td><em>Eimeria ahsata</em></td>
<td>MEO2015</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>Pending</td>
</tr>
<tr>
<td><em>Eimeria alabamensis</em></td>
<td>MEO2015</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>Pending</td>
</tr>
<tr>
<td><em>Eimeria tamiasciuri</em></td>
<td>MEO2015</td>
<td>Feecal Sample</td>
<td>mt genome</td>
<td>Pending</td>
</tr>
</tbody>
</table>
APPENDIX 6: DIVERGENT NUCLEAR 18S rDNA PARALOGS IN A TURKEY COCCIDUM, EIMERIA MELEAGRIMITIS, COMPLICATE MOLECULAR SYSTEMATICS AND IDENTIFICATION

(The contents of this chapter have been published as follows: Shiem El-sherry, Mosun E. Ogedengbe, Mian A. Hafeez, John R. Barta (2013). “Divergent nuclear 18S rDNA paralogs in Turkey coccidium, Eimeria meleagrimitis, complicate molecular systemics and identification”. International Journal for Parasitology 43: 679–685)
Divergent nuclear 18S rDNA paralogs in a turkey coccidium, *Eimeria meleagrimitis*, complicate molecular systematics and identification

Shiem El-Sherry, Mosun E. Ogedengbe, Mian A. Hafeez, John R. Barta

A R T I C L E   I N F O

Article history:
Received 21 December 2012
Received in revised form 19 March 2013
Accepted 20 March 2013
Available online 30 April 2013

Keywords:
Coccidiosis
Ribosomal DNA
Cytochrome c oxidase subunit I (COI)
DNA barcoding
Molecular taxonomy
Gene duplication
Diagnostics
Phylogeny

A B S T R A C T

Multiple 18S rDNA sequences were obtained from two single-oocyst-derived lines of each of *Eimeria meleagrimitis* and *Eimeria adenoeides*. After analysing the 15 new 18S rDNA sequences from two lines of *E. meleagrimitis* and 17 new sequences from two lines of *E. adenoeides*, there were clear indications that divergent, paralogous 18S rDNA copies existed within the nuclear genome of *E. meleagrimitis*. In contrast, mitochondrial cytochrome c oxidase subunit I (COI) partial sequences from all lines of a particular *Eimeria* sp. were identical and, in phylogenetic analyses, COI sequences clustered unambiguously in monophyletic and highly-supported clades specific to individual *Eimeria* sp. Phylogenetic analysis of the new 18S rDNA sequences from *E. meleagrimitis* showed that they formed two distinct clades: Type A with four new sequences; and Type B with nine new sequences; both Types A and B sequences were obtained from each of the single-oocyst-derived lines of *E. meleagrimitis*. Together these rDNA types formed a well-supported *E. meleagrimitis* clade. Types A and B 18S rDNA sequences from *E. meleagrimitis* had a mean sequence identity of only 97.4% whereas mean sequence identity within types was 99.1–99.3%. The observed intraspecific sequence divergence among *E. meleagrimitis* 18S rDNA sequence types was even higher (approximately 2.6%) than the interspecific sequence divergence present between some well-recognized species such as *Eimeria tenella* and *Eimeria necatrix* (1.1%). Our observations suggest that, unlike COI sequences, 18S rDNA sequences are not reliable molecular markers to be used alone for species identification with coccidia, although 18S rDNA sequences have clear utility for phylogenetic reconstruction of apicommplexan parasites at the genus and higher taxonomic ranks.

1. Introduction

Coccidiosis is a worldwide disease affecting turkeys and other galliform birds. Seven *Eimeria* spp. that infect turkeys have been described: *Eimeria adenoeides*; *Eimeria dispersa*; *Eimeria gallopavonis*; *Eimeria innocua*; *Eimeria meleagridis*; *Eimeria meleagrimitis*; and, *Eimeria subtunda* (see Chapman, 2008). Description of these species was based primarily on morphometric and biological characteristics. However, these parameters are insufficiently precise to reliably differentiate between species because morphometric and biological features frequently overlap and variations in these features were reported among strains and isolates in a single species (Long et al., 1977). Molecular data have become invaluable for the differentiation and classification of *Eimeria* spp.

The ssrRNA (18S rRNA) gene has been used extensively for classification of apicomplexan parasites (Morrison et al., 2004). Nucolar rDNA sequences have been used successfully in defining many taxonomic groups and proved to be good targets for species differentiation in the case of chicken *Eimeria* spp. (see Barta et al., 1997). Recently, however, two distinct types of 18S rDNA were reported for the chicken parasite, *Eimeria mitis*, in which sequence divergence between the two types was 1.3–1.7% whereas the sequence divergence within types (0.3–0.6%) was much more limited (Vrba et al., 2011). To put these levels of sequence divergence between the two rDNA types of *E. mitis* in perspective, the sequence variability between the rDNA sequences of *Eimeria tenella* and *Eimeria necatrix* is approximately 1.1%, less than the variability between rDNA types within a single *Eimeria* sp. Finding multiple, distinct rDNA sequences was previously reported for *Plasmodium* spp. (McCutchan et al., 1988; Nishimoto et al., 2008) in which up to three paralogous rDNA types may occur. However, Vrba et al. (2011) were the first to describe such divergent 18S rDNA loci within single oocyst lines of *Eimeria*.

The gene encoding mitochondrial cytochrome c oxidase subunit I (COI) has been widely used for phylogenetic analysis of many
organisms due to its near universal presence in organisms that use oxidative phosphorylation as an energy source (Hebert et al., 2003, 2004). Partial COI sequences have proven to be effective species-specific markers for Eimeria spp. and related coccidia (Ogedengbe et al., 2011).

In the present work, multiple single-oocyst-derived lines of the turkey parasites, E. adenoeides and E. meleagritmis, were generated and then both mitochondrial COI and nuclear 18S DNA genes were sequenced from each of the generated lines of parasites. Surprisingly high sequence diversity was found among nuclear 18S rRNA copies, suggesting that this genetic target may be unreliable as a species-specific marker; these observations are contrasted with the utility of the mitochondrial COI for the same purpose.

2. Materials and methods

2.1. Parasites – strains and generation of single-oocyst-derived lines

Eimeria meleagritmis USMN08-01 was isolated by Dr. H.D. Chapman (University of Arkansas, Fayetteville AR, USA) from a litter sample originating on a turkey farm in Minnesota, USA in 2008. Specific identity was established by demonstrating the presence of macrogamonts in histological sections of the jejunum and by the size of the oocysts (~20 by 17 µm) (H.D. Chapman, personal observation).

An isolate of E. adenoeides was originally obtained from a commercial turkey flock in Ontario, Canada, in approximately 1985. This isolate has been propagated periodically in specific parasite-free pouls in the Campus Animal Facility, University of Guelph, Guelph, Ontario, Canada (CAF), since that time. Oocyst dimensions (~18.7 by 14.2 µm) and sequence data (18S rRNA and mitochondrial COI partial sequences – data not shown) agree with the description of the KCH strain of E. adenoeides as defined by Poppelstein and Vrba (2011).

Freshly passed E. adenoeides KCH and E. meleagritmis USMN08-01 were prepared using single-oocyst isolations in vivo (four mono-specific lines for each species) and propagated in specific parasite-free pouls in CAF. Turkey pouls were provided feed (four mono-specific lines for each species) and propagated in a commercial turkey flock in Ontario, Canada, in approximately 1985. Turkey pouls were provided feed (four mono-specific lines for each species) and propagated in a commercial turkey flock in Ontario, Canada, in approximately 1985. Oocyst dimensions (~18.7 by 14.2 µm) and sequence data (18S rRNA and mitochondrial COI partial sequences – data not shown) agree with the description of the KCH strain of E. adenoeides as defined by Poppelstein and Vrba (2011).

Freshly passed E. adenoeides KCH and E. meleagritmis USMN08-01 were prepared using single-oocyst isolations in vivo (four mono-specific lines for each species) and propagated in specific parasite-free pouls in CAF. Turkey pouls were provided feed (four mono-specific lines for each species) and propagated in a commercial turkey flock in Ontario, Canada, in approximately 1985. Oocyst dimensions (~18.7 by 14.2 µm) and sequence data (18S rRNA and mitochondrial COI partial sequences – data not shown) agree with the description of the KCH strain of E. adenoeides as defined by Poppelstein and Vrba (2011).

Freshly passed E. adenoeides KCH and E. meleagritmis USMN08-01 were prepared using single-oocyst isolations in vivo (four mono-specific lines for each species) and propagated in specific parasite-free pouls in CAF. Turkey pouls were provided feed (four mono-specific lines for each species) and propagated in a commercial turkey flock in Ontario, Canada, in approximately 1985. Oocyst dimensions (~18.7 by 14.2 µm) and sequence data (18S rRNA and mitochondrial COI partial sequences – data not shown) agree with the description of the KCH strain of E. adenoeides as defined by Poppelstein and Vrba (2011).

Freshly passed E. adenoeides KCH and E. meleagritmis USMN08-01 were prepared using single-oocyst isolations in vivo (four mono-specific lines for each species) and propagated in specific parasite-free pouls in CAF. Turkey pouls were provided feed (four mono-specific lines for each species) and propagated in a commercial turkey flock in Ontario, Canada, in approximately 1985. Oocyst dimensions (~18.7 by 14.2 µm) and sequence data (18S rRNA and mitochondrial COI partial sequences – data not shown) agree with the description of the KCH strain of E. adenoeides as defined by Poppelstein and Vrba (2011).

2.2. PCR

Nuclear ssrDNA was amplified using universal eukaryote-specific primers (Medlin et al., 1988): Medlin A (5′-AACCTGGTATCC TGCCAGT-3′) and Medlin B (5′-GATCCCTGCAAGTTACCTAC-3′). The mitochondrial COI gene of E. adenoeides was amplified using primers 400F (5′-GDTTCAGGTRTGTTGAGAC-3′) and 1202R (5′-CAACRAHYGCAACAGATAA-3′). More degenerate primers 10F (5′-GWGDSWGGWRYGWTGGAC-3′) and 500R (5′-CATRTGIDGDCCAWAC-3′) were used to amplify a portion of the COI gene for E. meleagritmiss. PCRs were carried out in an MJ Mini thermal cycler (Bio Rad, CA, USA). PCRs contained ~100 ng of genomic DNA from each Eimeria sp., 50 mM MgCl2, 1 mM dNTPs, 10 × PCR buffer and 0.4 U Platinum Taq (Invitrogen). The PCR thermal profile was as follows: initial heat activation of polymerase at 96 °C for 10 min; 35 cycles of denaturation at 94 °C for 30 s, annealing at 48 °C for 30 s (for both 18S rRNA and COI reactions), extension at 72 °C for 1 min; and a final extension at 72 °C for 10 min. Both negative and positive template control reactions were included with each PCR run. PCR products were electrophoresed on a 1.5% agarose submarine gel in 1× Tris–Acetate–EDTA (TA) buffer at 109 V for 45 min. The resulting gel was stained with ethidium bromide and the size of products estimated by comparison with a 100 bp to 10 kb DNA ladder (Bio Basic, Inc., Mississauga, ON, Canada).

A QIA quick gel extraction kit (Qiagen, Toronto ON, Canada) was used to purify excised bands. The resulting DNA was then sequenced in both directions with the forward and reverse amplification primers using an ABI Prism 7000 Sequence Detection System (Applied Biosystems Inc., Foster City CA, USA) available in the Molecular Biology Unit of the Laboratory Services Division, University of Guelph (Guelph, ON, Canada). For the COI products, DNA sequencing was accomplished exclusively using PCR product templates. PCR-generated DNA template resulting from the Medlin A/B primer pair could not be successfully sequenced directly for any of the four E. adenoeides or E. meleagritmiss lines. Therefore, the 18S rDNA products from E. adenoeides KCH-Line 4 and KCH-Line 5 as well as E. meleagritmiss USMN-08-01 Line 4 and USMN-08-01 Line 5 were cloned individually using the TOPO TA cloning kit (Invitrogen). Recombinant clones from each of these four parasite lines were selected and the size of each cloned insert was confirmed using M13 Forward/M13 Reverse PCR amplification from putative positive colonies (data not shown). Clones containing inserts of approximately 1800 bp were cultured overnight in Luria broth (LB) and plasmid DNA isolated using Quick plasmid miniprep Kit (Invitrogen). The inserts in the resulting plasmids were sequenced completely in both directions using M13 Forward and M13 Reverse primers as well as four internal sequencing primers: 676F (5′-GGTTCAGTATGAAAGCGCTCGTA-3′); 1139R (5′-CATATCCTTTAAGGATGTC-3′); 1125F (5′-GAAACTTAAAGGAATTG-3′); and 696R (5′-TACGGAGCCTTYYAATGCAAC-3′).
2.3. Contig generation, sequence alignment and phylogenetic analyses

Chromatograms from sequencing reactions were used to form high quality contigs with the de novo assembler in Geneious version 6.1 created by Biomatters, available from http://www.geneious.com/; all PCR primer sites were identified in the resulting contigs. In addition to the COI and 18S rDNA sequences obtained in the present work, all 18S rDNA and COI sequences available for *E. adenoeides* and *E. meleagrimitis* as well as sequences from all available *Eimeria* spp. of galliform birds were retrieved from GenBank.

All publicly available and newly generated sequences were aligned using MUSCLE implemented within Geneious and then alignments were checked manually to maximize sequence similarities. After alignment, aligned sequences were trimmed to remove the forward and reverse amplification primer sites prior to phylogenetic analyses.

The resulting sequence alignments were analyzed phylogenetically using a number of tree-building methodologies, including Maximum Likelihood (ML), Maximum Parsimony (MP) and Bayesian inference (BI), all executed from within Geneious. For MP analyses, PAUP (V.4.10b, Swofford, 2003) was used with the parsimony criterion and equally weighted characters; *Cyclospora colobi* (AF111186) was used as the outgroup taxon for 100 bootstrap replicates. For BI, MrBayes (Huelsenbeck and Ronquist, 2001) was executed using a General Time Reversible (GTR) model and 6 gamma rate variation categories; *C. colobi* (AF111186) was again used as the outgroup taxon for 100 bootstrap replicates. For ML analyses, PhyML (Guindon and Gascuel, 2003) was used with the same GTR nucleotide substitution model as for the BI; 100 ML bootstrap replicates were performed on each dataset.

3. Results

3.1. PCR amplification and sequencing of the 18S rDNA from *E. adenoeides* and *E. meleagrimitis*

Medlin A and Medlin B universal primers successfully amplified ~1.8 kb fragments from all four single-oocyst derived lines of *E. adenoeides* KCH and four lines of *E. meleagrimitis* USMN-08-01. However, neither the Medlin A nor the Medlin B primers were able to provide high quality sequence over any distance from the priming site when gel-purified PCR products were used as template. For example, sequencing of the 1.8 kb PCR products from the *E. adenoeides* lines using the Medlin A primer generated high quality sequence for only ~80 nucleotides at which point the chromatograms became unreadable; these results were the first indication of the presence of multiple divergent copies of this gene in both species. Subsequent sequence analysis of clones (see below) demonstrated that the sequencing failure was the result of an indel at nucleotide 135 of the PCR product (see Fig. 1). Similar effects of indels on sequencing reactions from the PCR products were encountered with the 18S rDNA PCR products from all *E. meleagrimitis* lines. For this reason, the PCR products, generated using the Medlin A and Medlin B primer pair from two single-oocyst derived lines from each for *E. adenoeides* and *E. meleagrimitis*, were cloned and multiple 18S rDNA clones from two lines from both species were sequenced in both directions.

For *E. adenoeides*, 17 near-complete 18S rDNA clones were sequenced in both directions (eight from KCH-Line 4 (KC305169 to KC305176) and nine from KCH-Line 5 (KC305177 to KC305185)); that were 1791 to 1793 bp in length (including Medlin primers). For *E. meleagrimitis*, 15 clones were sequenced completely in both directions (five from USMN-08-01-Line 4 (KC305187 to KC305191) and 10 from USMN-08-01-Line 5 (KC305186 and KC305192 to KC305200)); with the lengths of sequences from 1795 to 1797 bp, including Medlin primers.

3.2. Analysis of intraspecific rDNA sequence variation

For *E. meleagrimitis*, near complete 18S rDNA sequences were aligned. Examination of the alignment revealed three apparently distinct types of 18S rDNA sequences: Type A consisted of four clones (one from USMN-08-01-Line 4; three from USMN-08-01-Line 5) with 99.3% identity across the entire sequence length and, on average, only six base differences (range 3–9 nucleotides) between any two members of this type; Type B contained nine clones (three from USMN-08-01-Line 4; six from USMN-08-01-Line 5) with 99.1% sequence identity and average within group variation of 5.8 base differences (range 0–11 differences); Type C consisted of two clones (one from USMN-08-01-Line 4; one from USMN-08-01-Line 5) that appeared to have high sequence identities with Type A sequences over a portion of their length and the

---

**Fig. 1.** Illustration of the effect of indels on the sequencing of 18S rDNA directly from PCR products. Sequences of four 18S rDNA clones and their corresponding GenBank accession numbers are aligned above portions of four separate sequencing chromatograms generated from PCR products and primed using the Medlin F primer. In each case, sequence quality and readability degraded immediately downstream of an indel (arrow). Such indel events in both *Eimeria meleagrimitis* and *Eimeria adenoeides* prevented 18S rDNA PCR products from being sequenced directly and necessitated cloning of PCR products into plasmids to permit full-length sequences to be obtained.
remains of the sequence had high sequence identity with the Type B sequences. Type C sequences may therefore represent PCR amplification hybrids. Mean pairwise sequence identities among Type A (99.7%) and Type B (99.6%) sequences were high; in contrast, sequence identity across all Types A and B sequences was only 96% and pairwise identities between sequences belonging to different types ranged from 96.8 to 97.4%, much lower than within-type pairwise identities. Types A and B sequences shared indels at positions 736 and 1371 and a single 18S rDNA clone from USMN-08-01-Line 5 had a unique indel at position 885.

For *E. adenoeides*, near complete 18S rDNA sequences formed a single, relatively homogenous, collection of 17 sequences (eight from KCH-Line 4 and nine from KCH-Line 5). All of these clones share 99.8% sequence identity with minor within group variation that averaged only 2.5 base differences between clones; among the 17 clones, five clones (three from KCH-Line 4 and two from KCH-Line 5) shared an indel at position 129 and a single clone from KCH-Line 4 demonstrated an indel at position 71.

### 3.3. Mitochondrial cytochrome c oxidase subunit I partial gene sequences

The 400F/1202R or 10F/500R primer sets successfully amplified and primed the sequencing of partial COI sequences from *E. adenoeides* (KC346357 to KC346360, 803 bp including primers) and *E. meleagrimitis* (KC346351 to KC346354, 513 bp including primers). All COI sequences from different single-oocyst derived lines of *E. adenoeides* were identical as were all similarly derived lines of *E. meleagrimitis*; there was no line to line variation observed in either species. Sequences of *E. adenoeides* shared identity with COI gene sequence of *E. adenoeides* strain KCH in GenBank (FR846202). There were no available COI sequences for *E. meleagrimitis* in GenBank.

### 3.4. Phylogenetic analyses using COI and 18S rDNA sequences

The 18S rDNA sequence dataset included 140 rDNA sequences in total with new 18S rDNA sequences for *E. meleagrimitis* and *E. adenoeides* from turkeys aligned with all publicly available sequences from *Eimeria* spp. of various galliform birds including: *Eimeria* spp. from chickens; *Eimeria* spp. of pheasants (*Phasianus colchicus*); and *Eimeria* spp. of quail (*Alectoria graeca*). An *Eimeria* sp. from a white-backed woodpecker (*Dendrocopos leucotos*) (FN298443) and two *Isospora* spp. (*Isospora Gryphonis* and *Isospora robini*) infecting non-galliform birds (*Carduelis tristis* and *Turdus migratorius*, respectively) were also included; a number of *Cyclospora* sp. formed the taxonomic outgroup for the phylogenetic analyses. All included sequences are identified in Fig. 2. Fifteen 18S rDNA sequences retrieved from GenBank were excluded from the analyses because they are likely hybrids that arose during PCR amplification of 18S rDNA from a mixed parasite species template (i.e. fecal samples from chickens or turkeys); the accession numbers of the excluded sequences follow: EU044768; EU044770; EU044771; EU044772; EU044774; EU044775; EU044776; EU044777; EU044779; EU044784; HM117011; HM117015; HM117016 and EF210326. Although there exists a single GenBank record for the 18S rDNA of *E. meleagrimitis*, this sequence is of relatively low quality judging by numerous unique indel events in regions of the rRNA gene that are highly conserved in all other *Eimeria* spp.; for this reason, sequence AF041437 was not included in this analysis. In addition, two newly generated sequences for *E. meleagrimitis* that were apparent Type A/Type B hybrids (accession numbers are KC305188 and KC305198) were also excluded from the analyses.

Phylogenetic analyses of aligned 18S rDNA sequences using ML, MP and Bayesian analysis resulted in trees with similar topologies and high bootstrap or nodal support values. The new 18S rDNA sequences from *E. meleagrimitis* clustered into two distinct clades (Type A with five sequences and Type B with 11 sequences) that together formed a well-supported *E. meleagrimitis* clade (Fig. 2). The two 18S rDNA types were each strongly supported (ML bootstrap = 100% and BI posterior probability of 1). In addition to the sequences obtained in the present study, there were three unnamed *Eimeria* spp. 18S rDNA sequences obtained from turkey fecal samples that were shown to belong to Type A (HM117012) or Type B (HM117013 and HM117014). For *E. adenoeides*, all new 18S rDNA sequences and two sequences obtained from GenBank formed one well-supported monophyletic clade that was the sister group to the *E. meleagrimitis* clade.

For COI sequences analysis, 92 sequences were used for the analyses with *G. Gryphonis* (KC346355) and an unnamed avian *Isospora* sp. (KC346356) acting as the functional outgroup. Ten sequences available from GenBank were excluded in the COI-based analysis because they were suspected to be PCR amplification hybrids; the accession numbers of the excluded sequences follow: FJ236431; FJ236432; FJ236440; FJ236441; HM771682; HM771683; HM771684; HQ221885; HQ702084 and AB564272. Sequences from *E. meleagrimitis* and *E. adenoeides* (including a previously unnamed *Eimeria* sp. of turkeys – HM117017 of Miska et al., 2010), each formed well supported monophyletic groups (Fig. 2) with markedly lower intraspecific genetic distances compared with the 18S rDNA tree. Sequences of *E. adenoeides* and *E. meleagrimitis* belonging to a clade that included unnamed *Eimeria* spp. from turkeys, pheasant and peafowl, *E. necatrix* and *E. tenella* from chickens. This large clade was the sister group to another large clade that contained the remaining five *Eimeria* spp. that infect chickens (*Eimeria acervulina, Eimeria brunetti, Eimeria maxima, E. mitis* and *Eimeria praecox*). These same relationships were also present in the 18S rDNA-based tree.

### 4. Discussion

The small subunit 18S rRNA gene is one of the most widely used genes in the study of phylogenetic relationships among species. rRNA gene sequences are easy to access due to the presence of highly
conserved flanking regions that help the universal primers to easily amplify and sequence this gene. This genetic locus has the added attraction of multiple genomic copies that made this region a good PCR target because the multiple copies increase sensitivity. For these reasons, the first large-scale phylogenetic studies depended on this gene (Field et al., 1988). Although the existence of multiple, sometimes many, copies of rDNA was known for eukaryotic genomes, any variations that arose among the copies within a genome were thought to be quickly homogenized through conversion and chromosomal crossover in a process known as concerted evolution (Hillis and Dixon, 1991). However, in some apicomplexan parasites it has been recognized that there may be up to three distinct rDNA types that coexist and persist within the nuclear genomes of species of Plasmodium (see McCutchan et al., 1988), Babesia (see Reddy et al., 1991) and Cryptosporidium (see Le Blancq et al., 1997).

Small subunit (18S) rRNA transcripts originating from distinct sets of ribosomal genes have been detected in several Plasmodium spp. (McCutchan et al., 1988). These transcripts are expressed differentially during the life cycle of the parasite (Gunderson et al., 1987; Li et al., 1997). This observation has been interpreted as an adaptation to the different host environments to which the parasite is exposed during its life cycle. One rRNA type has been shown to be expressed only during a particular growth stage of the organism and another type is expressed at a different stage, perhaps as an adaptation for immune-evasion in different hosts (Gunderson et al., 1987; Velichutina et al., 1998; Mercereau-Puijalon et al., 2002). It is now thought that a birth-and-death model under purifying selective pressure may best explain why the same rDNA type can be present in different parasite species while still permitting the appearance of new rDNA types through a variety of processes (Rooney 2004; Nei and Rooney, 2005). Bhoora et al. (2009) discovered extensive 18S rDNA sequence variation within Theileria equi and have characterized three genetically distinct types of rDNA from this parasite in South Africa. Although not as striking as for T. equi, the same authors described sequence variation within rDNAs of Babesia caballi as well.

Vrba et al. (2011) were the first to describe the presence of two types of nuclear 18S rDNA in an oocyst-cloned line of a coccidian parasite, E. mitis. They obtained six 18S rDNA sequences from five single-oocyst derived lines of E. mitis that grouped into two distinct types corresponding to published sequences submitted under the names Eimeria mivati and E. mitis; at the same time, the COI sequences from these same lines were all identical to one another and to the single E. mitis COI sequence available in GenBank. Vrba et al. (2011) concluded that E. mivati and E. mitis are the same species and that the appearance of two species using rDNA sequences as a marker was simply due to insufficient sampling. In the present study we discovered similar, but even more divergent, paralogous 18S DNA types in single-oocyst derived lines of the turkey parasite E. meleagrimitis. COI sequences obtained from multiple parasite lines confirmed that only a single parasite species, E. meleagrimitis, was present in each single-oocyst derived monospecific line and that these different rDNA types were both found in each line. The phylogenetic reconstructions showed that these two rDNA types form two distinct and well-supported clades. Previously reported 18S rDNA sequences from unnamed Eimeria spp. infecting turkeys (Miska et al., 2010) could be assigned to E. meleagrimitis 18S rDNA Type A (HM117012) or Type B (HM117013, HM117014) unambiguously. Eimeria spp. infecting galliform birds formed a well-supported clade that was the sister group to an Eimeria sp. found in a woodpecker. In the 18S rDNA-based tree, Eimeria spp. of partridge, pheasants and turkeys formed monophyletic groups but the Eimeria spp. of chickens were found in two monophyletic groups (five Eimeria spp. that infect the chicken intestine and two Eimeria spp. that infect the ceca of chickens). As suggested previously based on biological (Barta et al., 1997) and molecular

(Miska et al., 2010) data, the cecal parasites E. tenella and E. necatrix infecting chickens formed a well-supported monophyletic grouping found within a clade containing Eimeria spp. infecting turkeys, pheasants and peafowl, suggesting that these cecal parasites of chickens may have arisen as a result of a host switch event.

Sequence divergent between 18S rDNA Types A and B within E. meleagrimitis (97.4%) is higher than diversity found in E. mitis (98%), Cryptosporidium parvum (99%) or Babesia bigemina (99%). The intraspecific sequence divergence among E. meleagrimitis 18S rDNA sequence types was even higher (approximately 2.6%) than the interspecific sequence divergence present between some well-recognized species such as E. tenella and E. necatrix (1.1%). However, sequence divergence between 18S rDNA types in E. meleagrimitis was much lower than the diversity observed between 18S rDNA types in some Plasmodium spp. such as Plasmodium falciparum (89%, McCutchan et al., 1988), Plasmodium vivax (92%, Qari et al., 1994) or Plasmodium cynomolgi (90%, Corredor and Enea, 1994).

One complication that arose during the analysis of the 18S rDNA sequences that we obtained was the appearance of intermediate sequences (Type C) that appeared to belong clearly to neither Type A nor Type B rDNA sequences from E. meleagrimitis. Although it is possible that these intermediate forms represent actual rDNA repeats, we suspect that they are chimeric PCR products formed of fragments from both Types A and B; thus we excluded two of the newly obtained 18S sequences from E. meleagrimitis from the analyses. The frequency of chimeric molecules in PCR products has been estimated to range from 4% to 20% (Robison-Cox et al., 1995) when amplifications are made from mixed DNA templates that share PCR amplification primer sites. In the case of PCR amplification of 18S rDNA using the primers of Medlin et al. (1988), generation of such PCR artifacts (chimeric amplicons) can occur even from monospecific parasites as in the present study; such artificial sequences can suggest the presence of unique rDNA sequences that do not actually exist in nature. Our recognition of two apparent chimeric sequences among 15 18S rDNA clones highlights that such chimeric sequences are not at all uncommon and should be sought out and removed from analyses wherever possible. For Eimeria spp. and, potentially for other Apicomplexa as well, it should be noted that such chimeric sequences can arise even from a monospecific DNA template because, functionally, the DNA template from a species containing two or more distinct rDNA types is actually a mixed template. This observation suggests that multiple (-5–10) rDNA sequences should be obtained from even monospecific coccidian samples to guard against unwittingly sequencing either only one type of rDNA from a species when more than one type exist or a chimeric molecule that is not actually found within the target organism.

Checking for potentially chimeric sequences is relatively straightforward if there are a reasonable number of sequences from the taxon of interest. During data collection for our phylogenetic analyses, 15 18S rDNA and 10 COI sequences retrieved from GenBank were excluded because they were suspected to be hybrids between different species, in addition to the two apparent hybrids detected in our newly generated E. meleagrimitis 18S rDNA sequences. Detection of the hybrids required use of a basic local alignment search tool (BLAST) to search each suspect sequence separately against the GenBank database and then examining the sequence similarities to the BLAST search hits. Hybrids were frequently detected because they were ‘too unique’; they would hit one or more Eimeria sp. but the maximum sequence identity with other sequences ranged was frequently in the 97–98% range. When the suspect sequence was then pairwise aligned with the closest hits, the sequences were nearly identical over a portion of suspect sequence’s length but highly divergent with the remainder; searching the divergent region of the suspect sequence alone
would frequently show that this region was nearly identical to a second *Eimeria* sp. For example, one 18S rDNA sequence retrieved from GenBank (EU044771) when aligned with its closest BLAST hits had a 650 bp portion with 100% identity to *E. tenella* (EF210326) while the following 862 bp portion was 99.4% similar to *E. acervulina* (U67115). Such a molecule does not likely exist in nature and should be excluded from phylogenetic analyses.

The presence of two different 18S rDNA types in *E. meleagrimitis* raises the question as to why this apparent gene or locus duplication occurred and, more importantly, was retained in this species. The maintenance of multiple, developmentally regulated types of rDNA in *Plasmodium* spp. (Gunderson et al., 1987; Li et al., 1997) has been suggested to be related to the movement of the parasite into different temperature environments. Perhaps the occurrence of *E. meleagrimitis* in *Meleagris gallopavo*, a bird of the temperate climatic zones, may have selected for distinct rDNA loci that produced distinct rDNA types with different temperature optima; however, there is no evidence to date that the divergent 18S rDNA types observed in the present study are expressed differentially.

Our observations illustrate once again that 18S rDNA is not a suitable molecular marker to be used alone for species identification with coccidia, although the gene has clear utility for phylogenetic reconstruction of apicomplexan parasites at the genus level and at higher taxonomic ranks. If the sources of the 18S sequences from *E. meleagrimitis* obtained in the current study were not known, it would have been logical to conclude that there were two closely related but distinct *Eimeria* spp. present in the sample; only the COI sequence data and knowledge that the DNA came from single-ooyest cloned parasites prevented this interpretation of the data; the same situation occurred with *E. mitis* (see Vrba et al., 2011). In contrast to the confusion of the 18S rDNA based tree, the phylogenetic tree constructed from multiple COI sequences showed that COI sequences of *E. meleagrimitis* and *E. adenoeides* formed well-supported monophyletic clades distinct from other *Eimeria* spp. infecting galliform birds. The COI locus is thus far free of the gene duplication and divergence that has been demonstrated for nuclear rDNA. These observations confirm the utility of the mitochondrial COI locus as a species-level genetic marker for *Eimeria* spp. infecting turkeys as well as of other galliform birds (i.e. Ogedengbe et al., 2011; Vrba et al., 2011) and we encourage COI-based DNA barcoding of coccidia as part of the species description or re-descriptions of any new, or newly isolated, species of *Eimeria* or related parasites.

Acknowledgements

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) and from the Federal Economic Development Agency for Southern Ontario (FedDev) to J.R.B. and scholarship support from the Ministry of Agriculture, Food and Rural Affairs (OMAFRA) and from the Engineering Research Council of Canada (NSERC), the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) and from the Engineering Research Council of Canada (NSERC), the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) and from the Engineering Research Council of Canada (NSERC). Technical assistance of Julie Cobean is greatly appreciated.

References


Velichutina, I.V., Rogers, M.J., McCutchan, T.F., Liebman, S.W., 1998. Chimeric rRNAs containing the GTPase centers of the developmentally regulated ribosomal rRNAs of *Plasmidium falciparum* are functionally distinct. RNA 4, 594–602.

APPENDIX 7: BIOLOGICAL RE-DESCRIPTION OF A GENETICALLY-TYPED, SINGLE OOCYST LINE OF THE TURKEY COCCIDUM, *Eimeria meleagrimitis Tyzzer 1929*

(The contents of this chapter have been published as follows: Shiem El-sherry, Mosun E. Ogedengbe, Mian A. Hafeez, John R. Barta (2014). “Biological re-description of a genetically typed, single oocyst line of the turkey coccidium, *Eimeria meleagrimitis tyzzer 1929*”. Parasitology Research 113:1135–1146)
Biological re-description of a genetically typed, single oocyst line of the turkey coccidium, *Eimeria meleagrimitis* Tyzzer 1929

S. El-Sherry · T. Rathinam · M. A. Hafeez · M. E. Ogedengbe · H. D. Chapman · J. R. Barta

For the purpose of re-describing the *Eimeria* species that infect the turkey (*Meleagris gallopavo*) and to establish benchmark biological information linked to genetic markers for each species, a strain of *Eimeria meleagrimitis* Tyzzer 1929 was obtained from a litter sample from a turkey farm in Minnesota, USA in 2008. Multiple pure lines were derived by infecting turkey poults with a single oocyst; one of these lines was then used to re-describe biological and morphological features of *E. meleagrimitis* in the turkey and to designate a neotype of *E. meleagrimitis* in the turkey. Oocyst morphometrics of this line matched those of this species as originally described by Tyzzer (Am J Hyg 10:269–383, 1929). Three asexual generations of merogony (the first generation of meronts large in size and the second and third generations small) were detected in the intestines before the onset of gametogony; no developmental stages were detected in the cecal pouches. No mortality was induced by this line of *E. meleagrimitis* even when turkey poults were infected with high doses of oocysts (up to $5 \times 10^5$ oocysts/bird) and despite the ability of *E. meleagrimitis* to induce severe mucosal damage in the upper and middle duodenum. Macroscopic lesions were characterized to provide a graded lesion scoring guide that should assist assessment of the severity of infections with this species in infected turkeys. The pathogenicity of the strain was investigated, and a significant reduction in weight gain and feed conversion ratio was observed with doses of $10^4$ oocysts/bird or more. The maximum yield of oocysts in the feces was obtained when birds were inoculated with $5 \times 10^3$ oocysts.

Introduction

Coccidiosis is a common disease of poultry, including turkeys (Long and Millard 1977), caused by protozoan parasites of the genus *Eimeria* that develop in epithelial cells of the intestine (Chapman 2008). The disease causes malabsorption of nutrients, reduced weight gain, and poor feed conversion and may lead to severe economic losses, especially in intensive production systems (Long 1973; McDougald and Fitz-Coy 2003). Whereas the life cycles and biology of *Eimeria* spp. from the chicken have been studied in great detail, this is not the case for those from the turkey (Chapman 2008). Seven *Eimeria* spp. have been described from this host, and of these, three species (*Eimeria adenoeides*, *Eimeria meleagrimitis*, and *Eimeria gallopavonis*) are considered to be highly pathogenic (McDougald and Fitz-Coy 2003). In this study, we provide further information on one of these species, *E. meleagrimitis*.

*E. meleagrimitis* is a common parasite of the turkey that is frequently identified in litter samples obtained from turkey farms (Chapman and Rathinam 2007; Jeffers and Bentley 1980). This species develops in the anterior and mid-intestine and can cause enteritis, hemorrhage, and extensive sloughing of the epithelium (Clarkson 1959; Joyner 1978). Despite its widespread occurrence in commercial turkey production, surprisingly little information exists regarding its life cycle and biology. *E. meleagrimitis* was first described in 1929 by Tyzzer and subsequently investigated by Hawkins, but it was not until 1959 that Clarkson provided a detailed description of this parasite (Clarkson 1959; Hawkins 1952; Tyzzer 1929). Subsequent investigations were carried out by Long...
et al. (1977) and Ruff et al. (1980). Some aspects of the biology and life cycle of *E. meleagrimitis* provided in these publications are contradictory, and most, if not all, of the original strains used by these workers have been lost, including that from the first description by Tyzzer, making it impossible to re-evaluate the species description or to compare recent isolates with a reference strain (Chapman 2008). Reliable diagnostic characteristics, including more detailed information regarding endogenous development, pathogenicity, and fecundity of this parasite, are essential if accurate methods for species identification and improved methods of control are to be developed. In the present study, we utilized a molecularly characterized, single-oocyst-derived line of *E. meleagrimitis*, in order to establish benchmark biological and morphometric features. We provide a formal description of *E. meleagrimitis* which may be of value to others in future investigations of this economically important parasite.

**Materials and methods**

**Origins and derivation of single-oocyst-derived lines of *E. meleagrimitis***

A strain of *E. meleagrimitis* was obtained from a litter sample received from a turkey farm near Willmar, MN, USA on February 26, 2008. A pure infection was derived by isolating single oocysts using the method described by Shirley and Harvey (1996) and subsequently propagating these in uninfected turkeys reared in an isolation unit at the coccidiosis facility, Department of Poultry Science, University of Arkansas, Fayetteville, AR, USA. This parent strain is referred to as *E. meleagrimitis* USMN08-01 and was propagated in specific parasite free poults in the animal isolation facility at the University of Guelph, Guelph, ON, Canada. Turkey poults were provided feed and water ad libitum; all experimental procedures were reviewed and approved by the University of Guelph’s Animal Care Committee and complied with the Canadian Council on Animal Care’s Guide to the Care and Use of Experimental Animals (second edition). A number of single-oocyst lines (USMN08-01-Line 1, USMN08-01-Line 4, USMN08-01-Line 5, and USMN08-01-Line 8) were derived as described by Remmler and McGregor (1964) with the modification that agar plugs carrying a single oocyst were given orally to the birds within gelatin capsules as described by El-Sherry et al. (2013). One of the lines (USMN08-01-Line 5) was used for all experiments described herein. The sequence of the mitochondrial cytchrome c oxidase subunit I and nuclear small subunit 18S rDNA loci of this single-oocyst-derived line of *E. meleagrimitis* was determined by El-Sherry et al. (2013).

Feces were collected from days 6 to 8 after oral inoculation of pouls with 1.5×10⁷ sporulated oocysts of *E. meleagrimitis* USMN08-01-Line 5. Oocysts were separated from the feces by salt flotation and, following suspension in 2.5 % potassium dichromate solution, sporulated by incubation at 26 °C on a rotary shaker for a minimum of 72 h (Reid and Long 1979). Oocysts were then collected by centrifugation (1,000×g for 10 min), the potassium dichromate solution decanted, and the resulting pellet re-suspended in 1× PBS. Oocysts were stored in a refrigerator at 4 °C until required for experiments. Sporulated oocysts were counted and diluted to the desired doses as described by Shirley (1995). Oocysts were stored for less than a month before experimental use.

**Oocyst measurements**

Measurements of the oocysts and all life cycle stages were made using a Provis AX70 photomicroscope (Olympus Canada, Richmond Hill, ON, Canada) fitted with a digital imaging device (Infinity3-1C, Lumenera Corporation, Ottawa, ON, Canada) controlled using iSolution Lite image analysis software (Hoskin Scientific, Burlington, ON, Canada). All measurements are reported as means followed by the range and sample size in parentheses.

**Characterization of the endogenous life cycle stages**

Twelve 18-day-old poults (Hybrid Turkeys, Kitchener, ON, Canada) were each infected with a single dose of oocysts and necropsied from 32 to 120 h post-inoculation (HPI). To ensure that sufficient numbers of parasites were present to permit their visualization in histological sections, birds necropsied to reveal early endogenous developmental stages of the life cycle were inoculated with larger numbers of oocysts than those being examined for later gametocyte development (Table 1). For each bird, the intestinal tract (IT) was removed, and 1-cm-long pieces were taken from eight locations as follows: (1) middle of the descending duodenal loop, (2) middle of the ascending duodenal loop, (3) jejunum approximately 3 cm proximal to Meckel’s diverticulum, (4) jejunum approximately 3 cm distal to Meckel’s diverticulum, (5) midpoint of the ileum (halfway between Meckel’s diverticulum and the ileocecal junction), (6) proximal 1 cm of the neck of the cecum, (7) middle of the cecal pouch, and (8) middle of the rectum. Samples were rinsed briefly with saline and then placed in freshly prepared Serra fixative solution [100 % ethanol (60 %, v/v), 37 % formaldehyde (30 %, v/v), and glacial acetic acid (10 %, v/v)] for 24 h at room temperature. Samples were then embedded in wax, sectioned at a thickness of ~4 μm, mounted on a glass slide, and then stained with hematoxylin and eosin (Animal Health Laboratory, University of Guelph). The resulting histological sections were examined for the presence of endogenous stages of the parasite life cycle. Measurements of each life cycle stage...
observed \((n=20)\) were made using the microscope and software as described previously.

Pathogenicity: macroscopic lesion scores

Forty-two 21-day-old poults were orally inoculated, in groups of four to five birds each, with \(4 \times 10^4, 1 \times 10^5, 1.5 \times 10^5, 2.5 \times 10^5\), or \(5 \times 10^5\) oocysts/poult. Two poults were sham-inoculated with water (uninfected controls). Poults were killed humanely 5 and 6 days post-inoculation and the IT examined for macroscopic lesions; sham-inoculated birds were necropsied at day 6 post-inoculation. Lesions were scored from 0 to 4 based on severity of the observed lesions in four regions of the IT (duodenum, jejunum, ileum, and ceca).

Pathogenicity: weight gain-to-feed conversion ratio

Ninety-six female Nicholas poults were raised until 13 days of age in a brooder in an isolation building. They were allocated to four treatments each comprising four cages of six poults (24 poults/treatment). Poults in treatments 1, 2, and 3 were inoculated with \(1 \times 10^4, 5 \times 10^4,\) or \(1 \times 10^5\) oocysts/poult, respectively. Poults in treatment 4 were sham-inoculated with water (uninfected controls). Body weights of all poults were obtained on the day of inoculation (day 0) and again on day 6 post-inoculation. Feed consumed from day 0 to 6 was determined and feed conversion calculated. The intestines of any birds that died were examined to determine if coccidiosis was the cause of death, and the percent of mortality attributable to coccidiosis was calculated for each inoculation dose.

Oocyst inoculum required to achieve maximal oocyst production

Thirty 3-week-old Hybrid turkey poults were randomly allocated to six cages (five birds per cage). Poults in each cage were inoculated with either \(1 \times 10^2, 10 \times 10^3, 5 \times 10^3, 10 \times 10^4, 2.5 \times 10^4,\) or \(5 \times 10^5\) oocysts. Fecal collection was accomplished on days 4–9 following infection by placing aluminum foil beneath the cage floors. The foil was removed each day, the feces collected, and the number of oocysts present determined. Fecundity was calculated by dividing the total oocyst production per bird by the dose of oocysts administered.

Data analysis of weight gains and feed conversions

All the data were expressed as mean±SEM and analyzed using SAS® Enterprise Guide 4.2® (SAS Institute Inc, Cary, NC, USA). Statistical analysis of weight gain, feed intake, and feed conversion was carried out by one-way ANOVA, and the means were separated using Student–Newman–Keuls test. Results with a \(p\) value ≤0.05 were considered as significantly different.

Results

Species re-description

**Taxonomic summary**

Apicomplexa
Conoidasida
Coccidia
Eucoccidiorida
Eimeriorina
Eimeriidae

*E. meleagrimitis* Tyzzer 1929

**Geographic origin and date of isolation of parent culture:** Near Willmar, MN, USA (near 45.117, −95.050), February 26, 2008.

**Type—host:** *Meleagris gallopavo* (Aves, Galliformes, Phasianidae, Meleagridinae).

**Type—locality:** Unknown, presumed near Cambridge, MA, USA or environs, near 42.337, −71.103 (Tyzzer 1929)

**Other localities:** MN, USA; Guelph, ON, Canada; Buxted, Llanelly, Thirsk, UK (see Joyner and Norton 1972); and elsewhere; likely cosmopolitan in turkey flocks

Description of sporulated oocysts

Oocysts were subspherical in shape with a doubled-contoured appearance to the oocyst wall (Fig. 1a). Sporulated oocysts measured \(18.9±1.6 \mu m\) (17–22)×\(15.7±1.4 \mu m\) (13–19, \(n=30\)) with a shape index (SI) of \(1.20±0.1\). A single refractile granule was typically detected in sporulated oocysts. Sporocysts measured \(10.9±0.4 \mu m\) (10–12)×\(6.1±0.7 \mu m\) (5–7, \(n=30\)) (Fig. 1b). Sporozoites possess anterior and larger posterior refractile bodies (Fig. 1b). A scatter plot of length
and width for the measured oocysts and sporocysts demonstrated the presence of a morphometrically homogeneous population (Fig. 2). The sporulation of the oocyst is exogenous, experimentally 48 h at 26 °C; length of sporulation at other temperatures is not known.

Endogenous development

The life cycle stages of *E. meleagrimitis* that were observed microscopically and their locations throughout endogenous development are summarized in Table 2, and the locations of asexual and sexual stages are illustrated diagrammatically in a drawing of the intestine (Fig. 3a, b.) respectively. No parasites were detected in the cecal pouch (i.e., distal to the cecal neck) at any sampled time post-inoculation.

**First asexual stage** At 32 HPI, sporozoites, trophozoites, and immature meronts were found within crypt cells mainly in the jejunum and to a lesser extent in the duodenum. The growth of the first-generation meronts below the host cell nuclei appeared to distort the cell shape (Fig. 4a); infected enterocytes looked swollen and their nuclei were displaced toward the lumen. The earliest mature first-generation meronts were detected at 40 HPI and appeared to consist of 45–60 merozoites. Mature meronts measured 13 μm (12–16)×10 μm (8–13, n=20). Immature first-generation meronts were seen in the lower part of the intestine (ileum, cecal neck, and rectum) at 40 HPI and matured at those locations. First-generation meronts were detected until 56 HPI.

**Second asexual stage** The earliest second-generation meronts were detected at 48 HPI in the jejunum principally and to a lesser extent in the duodenum. Few were found in the lower IT (ileum, cecal neck, and rectum). The second-generation meronts were located in the crypts and enterocytes lining the base of villi. Mature second-generation meronts measuring 5 μm (4–6)×5 μm (4–6) were numerous at 56 HPI. Mature meronts were located both above and below host cell nuclei and contained about eight to 12 merozoites. Maturing second-generation meronts continued to be observed until 64 HPI. No obvious pathologic alterations to the mucosa were associated with the first two asexual generations (Fig. 4b).

**Third asexual stage** Massive numbers of trophozoites initiating the third asexual generation were detected at 72 HPI infecting enterocytes lining the sides and tips of the villi throughout the IT (except the cecum) with the highest density of parasites observed in the jejunum. Up to three trophozoites could be detected in one host cell below or above the nucleus. Most mature third-generation meronts, containing about 10–12 merozoites, were observed at 96 HPI; each measured 6 μm (4–7)×6 μm (4–7). At this time, pathologic changes to the infected epithelium were evident in the jejunum. Mature meronts were detected throughout the IT except the cecum. Mature third-generation meronts were observed until 104 HPI (Fig. 4c).

**Gamont stage** Immature gamonts were detected at 112 HPI throughout the IT except the cecal pouch; the proximal cecal
neck was affected. Gamonts were observed within enterocytes lining the sides and tips of the villi but were not observed within the crypt epithelium. Gamonts in the upper and mid-intestine appear to be larger and more differentiated than in the lower regions at this time. No oocysts were detected at 112 HPI. Oocysts were observed at 120 HPI (pre-patent period 5 days), and numerous mature gamonts were still present. Female gamonts outnumbered male gamonts by a ratio of approximately 8:1. Mature female gamonts measured 15 μm (13–17)×12.0 μm (11–14), and mature male gamonts measured 14 μm (12–18)×11.1 μm (9–13) (Fig. 4d).

**Table 2** Location and density of *E. meleagrimitis* life cycle stages in histological sections

<table>
<thead>
<tr>
<th>Time post-inoculation (h)</th>
<th>Region of digestive tract</th>
</tr>
</thead>
<tbody>
<tr>
<td>DD</td>
<td>AD</td>
</tr>
<tr>
<td>32</td>
<td>+</td>
</tr>
<tr>
<td>40</td>
<td>+</td>
</tr>
<tr>
<td>48</td>
<td>+</td>
</tr>
<tr>
<td>56</td>
<td>+</td>
</tr>
<tr>
<td>64</td>
<td>+</td>
</tr>
<tr>
<td>72</td>
<td>++</td>
</tr>
<tr>
<td>80</td>
<td>++</td>
</tr>
<tr>
<td>88</td>
<td>++</td>
</tr>
<tr>
<td>96</td>
<td>++</td>
</tr>
<tr>
<td>104</td>
<td>++</td>
</tr>
<tr>
<td>112</td>
<td>++</td>
</tr>
<tr>
<td>120</td>
<td>++</td>
</tr>
</tbody>
</table>

+, ++, and +++ indicate an increasing density of parasites in different regions of the IT. DD descending duodenum, AD ascending duodenum, JBM jejunum before Meckel’s diverticulum, JAM jejunum after Meckel’s diverticulum, MI middle of ileum, CN cecal neck, MC middle of the cecum, MR middle of rectum.

**Fig. 3** Diagram illustrating the distribution of life cycle stages of *E. meleagrimitis* within an infected turkey: a asexual stages mainly in the upper and middle part of IT (duodenum and jejunum) with the highest density of the parasites present in jejunum; b sexual stages detected throughout the IT with the higher density in the lower part of the IT including cecal neck but not within the cecal pouches.
Pathogenicity

Bodyweight gain, feed intake, feed conversion, mortality, and macroscopic lesion score data of poults are presented in Tables 3 and 4. Significant reductions in weight gain from days 0 to 6 following infection of poults were observed compared with sham-infected poults. Higher doses ($5 \times 10^4$ or $10^5$ oocysts) resulted in significantly lower weight gains than poults infected with $1 \times 10^4$ oocysts suggesting that the growth suppression was dose-dependent. Infected poults had significant reductions in feed intake compared with uninfected poults over a 6-day period following inoculation, but there was no significant difference in feed intake between the heavily infected poults ($5 \times 10^4$ or $10^5$ oocysts/bird). Poults given $10^4$ oocysts consumed more food than poults given $10^5$ oocysts. With the exception of poults given $10^4$ oocysts, infected poults had significantly higher feed conversion ratios than uninfected poults. No lesions were found in sham-inoculated controls. No mortalities due to coccidial infection were observed with any of the infective doses.

Pathological lesions

Macroscopic lesions were necrosis, sloughing of the epithelium, ulceration, petechiae, and mucoid casts mainly in the duodenum and jejunum observed in heavily infected turkeys.

Table 3  Body weight gains, feed intake, and feed conversion of poults infected with E. meleagrimitis at various doses; control birds were sham-inoculated with saline

<table>
<thead>
<tr>
<th>Infection status</th>
<th>Oocyst dose$^a$</th>
<th>Weight gain (g)$^b$</th>
<th>Feed intake (g)$^c$</th>
<th>Feed conversion$^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uninfected</td>
<td>0</td>
<td>237±6 a</td>
<td>342±5 a</td>
<td>1.45±0.06 b</td>
</tr>
<tr>
<td>Infected</td>
<td>$10^4$</td>
<td>141±5 b</td>
<td>277±4 b</td>
<td>1.96±0.05 b</td>
</tr>
<tr>
<td></td>
<td>$5 \times 10^4$</td>
<td>96±9 c</td>
<td>254±9 c</td>
<td>2.84±0.43 a</td>
</tr>
<tr>
<td></td>
<td>$10^5$</td>
<td>92±4 c</td>
<td>248±3 c</td>
<td>2.70±0.12 a</td>
</tr>
</tbody>
</table>

Values in columns with no common lowercase letter differ significantly ($P \leq 0.05$). Each observation is the mean for four cages each containing six poults

$^a$ Number of oocysts administered to each poult

$^b$ Mean bodyweight gain±SEM from day 0 to 6 post-inoculation

$^c$ Mean feed consumed per poult from day 0 to 6 post-inoculation

$^d$ Feed conversion was calculated by dividing feed intake by bodyweight gain for each bird
Increasing the number of inoculated oocysts (4×10⁴, 1×10⁵, or 1.5×10⁵/poult) led to increased severity of pathogenic changes and associated higher lesion scores (Table 4). Further increases in inoculation dose (2×10⁵ or 5×10⁵ oocysts/poult) resulted in decreased macroscopically visible mucosal damage and correspondingly lower assigned lesion scores in the intestinal tract. This phenomenon was repeated in birds necropsied at 6 days post-infection. Increasing the infective dose not only affected the intensity of the lesion but it also altered the location of the resulting visible lesions. Once the challenge dose increased beyond a comparatively modest level, lesions became less intense (lower lesion scores) in the typical areas of infection (duodenum and jejunum), but lesions began to be seen in the lower reaches of the intestinal tract, including the ileum and cecal neck, that were not usually affected during infections initiated with fewer oocysts.

Reproductive potential and dose-dependent fecundity

The relationship between various inoculation doses and resultant oocyst shedding by a single poult is illustrated (Fig. 6a). The greatest total output of ~7.4×10⁷ oocysts/bird was obtained from birds inoculated with 5×10⁵ oocysts; poult inoculated with higher (1×10⁶ oocysts) or lower (1×10⁴ oocysts) doses shed slightly fewer total oocysts suggesting that a dose of approximately 5×10⁵ should generate the maximum number of total oocysts from turkey poult. The yield decreased with increasing inoculum doses (10⁴ to 5×10⁵ oocysts). Consequently, fecundity (oocysts produced per oocyst inoculated) decreased dramatically with increasing numbers of oocysts in the inoculum (Fig. 6b) from a peak of 558,238 to a low of 881 when the inoculating dose was 10⁵ and 5×10⁴ oocysts, respectively.

Reference molecular data

Molecular data specifically related to *E. meleagrimitis* USMN08-01-Line 5 as well as additional single-oocyst-derived lines of *E. meleagrimitis* have been published previously (El-Sherry et al. 2013). The sequence data specifically associated with *E. meleagrimitis* USMN08-01-Line 5 for which the biological re-description is provided herein are as follows:

Mitochondrial cytochrome c oxidase subunit 1 (COI) partial sequence. KC346353 El-Sherry et al. (2013) amplified and sequenced a portion of the mitochondrial COI gene from *E. meleagrimitis* USMN08-01-Line 5. Nuclear 18S rDNA near-complete sequences KC305186 and KC305192 to KC305200. Ten nuclear 18S rDNA sequences were obtained from *E. meleagrimitis* USMN08-01-Line 5 (see El-Sherry et al. 2013). Despite the wide intraspecific variation observed among the rDNA sequences, these sequences, individually or collectively, should be considered valid 18S rDNA genotypes for *E. meleagrimitis*.

Designation of a neotype and type deposition

A name-bearing type for *E. meleagrimitis* was never designated by Tyzzer (1929) nor have any specimens referable to his original species description been discovered subsequently. In an effort to stabilize the taxonomic status of *E. meleagrimitis* Tyzzer 1929, a neotype has been designated for *E. meleagrimitis* USMN08-01-Line 5 to act as the name bearing type for the species, and this neotype has been deposited to the Canadian Museum of Nature under catalogue

### Table 4 Lesion scores of poults infected experimentally with various doses of sporulated oocysts and then necropsied at 5 or 6 days post-inoculation with *E. meleagrimitis*

<table>
<thead>
<tr>
<th>Dose oocysts/bird</th>
<th>Number of birds</th>
<th>Duodenum</th>
<th>Jejunum</th>
<th>Ileum</th>
<th>Cecum</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 days post-inoculation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4×10⁴</td>
<td>5</td>
<td>1±0.7</td>
<td>1.4±0.8</td>
<td>0.8±0.4</td>
<td>0.2±0.4</td>
</tr>
<tr>
<td>1×10⁵</td>
<td>4</td>
<td>1.5±1.3</td>
<td>2±0.8</td>
<td>1</td>
<td>0.25±0.5</td>
</tr>
<tr>
<td>1.5×10⁵</td>
<td>4</td>
<td>1.75±0.5</td>
<td>2.25±1</td>
<td>1.5±0.6</td>
<td>0.5±0.6</td>
</tr>
<tr>
<td>2×10⁵</td>
<td>4</td>
<td>0.5±0.6</td>
<td>1.5±0.6</td>
<td>1.25±1</td>
<td>0</td>
</tr>
<tr>
<td>5×10⁵</td>
<td>4</td>
<td>1.5±0.5</td>
<td>2±0.8</td>
<td>2±0.8</td>
<td>0.5±0.6</td>
</tr>
<tr>
<td>6 days post-inoculation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4×10⁴</td>
<td>3</td>
<td>1.7±1.5</td>
<td>1.3±0.6</td>
<td>0.7±0.6</td>
<td>0</td>
</tr>
<tr>
<td>1×10⁵</td>
<td>4</td>
<td>3.3±0.5</td>
<td>1.5±0.6</td>
<td>1±0</td>
<td>1±0</td>
</tr>
<tr>
<td>1.5×10⁵</td>
<td>4</td>
<td>2.3±1</td>
<td>1.5±0.6</td>
<td>1.25±0.5</td>
<td>1.25±0.5</td>
</tr>
<tr>
<td>2×10⁵</td>
<td>3</td>
<td>2.7±0.6</td>
<td>1.3±0.6</td>
<td>1.3±0.6</td>
<td>1.3±0.6</td>
</tr>
<tr>
<td>5×10⁵</td>
<td>4</td>
<td>1.5±0.6</td>
<td>1.5±0.6</td>
<td>1.5±0.6</td>
<td>1.5±0.6</td>
</tr>
<tr>
<td>Sham inoculated controls</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Feature</td>
<td>Tyzzer (1929)</td>
<td>Hawkins (1952)</td>
<td>Clarkson (1959)</td>
<td>USMN08-01-line 5 (present study)</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>----------------------------------</td>
<td></td>
</tr>
<tr>
<td>Oocyst size</td>
<td>18 × 15.25 μm</td>
<td>19.17 × 16.28 μm</td>
<td>20.1 × 17.3 μm</td>
<td>18.9 × 15.7 μm</td>
<td></td>
</tr>
<tr>
<td>Prepatent period</td>
<td>144 h</td>
<td>144 h</td>
<td>116 h</td>
<td>120 h</td>
<td></td>
</tr>
<tr>
<td>Mature first generation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>6–9.5 × 5.8–7.7 μm</td>
<td>8–9 μm in diameter</td>
<td>17 × 13 μm</td>
<td>13 × 10 μm</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>Mainly in the lower part of the small intestine</td>
<td>–</td>
<td>Mainly jejunum less extent duodenum</td>
<td>Throughout IT except the cecal pouch; mainly duodenum and jejunum</td>
<td></td>
</tr>
<tr>
<td>Histological location</td>
<td>Deep in epithelium below host cell nuclei</td>
<td>–</td>
<td>In deep glands under host cells nuclei</td>
<td>In deep glands under host cells nuclei</td>
<td></td>
</tr>
<tr>
<td>No. of merozoites</td>
<td>16</td>
<td>9–10</td>
<td>80–100</td>
<td>45–60</td>
<td></td>
</tr>
<tr>
<td>First observed</td>
<td>–</td>
<td>66 h</td>
<td>48 h</td>
<td>40 h</td>
<td></td>
</tr>
<tr>
<td>Mature second generation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>–</td>
<td>12–14 μm in diameter</td>
<td>8 × 7 μm</td>
<td>5 × 5 μm</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>–</td>
<td>–</td>
<td>Mainly jejunum, less extent duodenum</td>
<td>Throughout IT except the cecal pouch; mainly duodenum and jejunum</td>
<td></td>
</tr>
<tr>
<td>Histological location</td>
<td>–</td>
<td>Tips of the villi, not in deep glands</td>
<td>In deep glands and spread up the villi</td>
<td>In the crypts and enterocytes lining the base of villi</td>
<td></td>
</tr>
<tr>
<td>No. of merozoites</td>
<td>–</td>
<td>10–12</td>
<td>8–16</td>
<td>8–12</td>
<td></td>
</tr>
<tr>
<td>First observed</td>
<td>–</td>
<td>96 to 108 HPI</td>
<td>66 HPI</td>
<td>56 HPI</td>
<td></td>
</tr>
<tr>
<td>Mature third generation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>–</td>
<td>–</td>
<td>8 × 7 μm</td>
<td>6 × 6 μm</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>–</td>
<td>–</td>
<td>Tend to spread down throughout the IT and the cecum</td>
<td>Throughout the IT except the cecal pouch</td>
<td></td>
</tr>
<tr>
<td>Histological location</td>
<td>–</td>
<td>–</td>
<td>Alongside the villi to the tips, Not in deep glands</td>
<td>Enterocytes lining the sides and tips of the villi. Not in deep glands</td>
<td></td>
</tr>
<tr>
<td>No. of merozoites</td>
<td>–</td>
<td>–</td>
<td>8–16</td>
<td>10–12</td>
<td></td>
</tr>
<tr>
<td>First observed</td>
<td>–</td>
<td>–</td>
<td>96 HPI</td>
<td>96 HPI</td>
<td></td>
</tr>
<tr>
<td>Mature gamonts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>–</td>
<td>7–8 μm in diameter</td>
<td>15 × 11 μm</td>
<td>Female 15 × 12 μm, male 14 × 11 μm</td>
<td></td>
</tr>
<tr>
<td>Intestinal location</td>
<td>–</td>
<td>Middle half of small intestine</td>
<td>Tend to spread down the IT and the cecum</td>
<td>Throughout the IT except the cecal pouch</td>
<td></td>
</tr>
<tr>
<td>Histological location</td>
<td>–</td>
<td>–</td>
<td>Epithelium cells particularly at the tips of the villi</td>
<td>Enterocytes lining the sides and tips of the villi. Not in deep glands</td>
<td></td>
</tr>
<tr>
<td>First observed</td>
<td>–</td>
<td>120 HPI</td>
<td>114 HPI</td>
<td>112 HPI</td>
<td></td>
</tr>
</tbody>
</table>

IT intestinal tract
number CMNPA 2013-0003. The neotype consists of a hapantotype series of four histological preparations of the intestinal tract of experimentally infected turkeys and a phototype of sporulated oocysts as follows: CMNPA 2013-0003.1—48 HPI demonstrating mature first-generation meronts; CMNPA 2013-0003.2—56 HPI demonstrating mature second-generation meronts; CMNPA 2013-0003.3—96 HPI demonstrating third-mature generation meronts; CMNPA 2013-0003.4—120 HPI demonstrating mature macrogamonts, microgamonts, and unsporulated oocysts; and CMNPA 2013-0003.5—a phototype of the sporulated oocyst of this parasite.

Comments
Some of the other Eimeria species infecting turkeys can be confused with E. meleagrimitis. For example, Eimeria dispersa develops in the duodenum and mid-intestine, but unlike E. meleagrimitis, E. dispersa is confined mainly to the upper intestine (duodenum) and spreads to the lower intestine during late merogony and gametogony (Doran 1978; Long and Millard 1979; Chapman 2008). Oocysts of E. dispersa are larger (22.7×18.8 μm, Tyzzer 1929; 23.8×19.8 μm; El-Sherry, unpublished observations) than those of E. meleagrimitis, have single-layered oocyst walls, and lack the characteristic single polar granule typical of E. meleagrimitis. Other species of Eimeria affecting the upper and mid-intestine include Eimeria innocua (Moore and Brown 1952) and Eimeria subrotunda (Moore et al. 1954), but no polar granules were reported from the oocysts of either of these species. Matsler and Chapman (2006) re-described Eimeria meleagrisidis and documented that E. meleagrisidis has larger oocysts (24.9×17 μm) and affects mainly the middle and lower part of the intestine; further, almost all endogenous life cycle stages can be found infecting the middle and distal reaches of the cecum. Of the remaining Eimeria spp. infecting turkeys, oocysts of E. gallopavonis are much larger (27.1×17.2 μm, Hawkins 1952), and its endogenous development is restricted to the lower part of the intestine; likewise, E. adenoideis is restricted to the lower part of the intestinal tract and has been reported to possess larger oocysts (e.g., 25.6×16.25 μm, Moore and Brown 1951). The only oocysts shed by Eimeria-infected turkeys that could be confused morphologically with E. meleagrimitis is the “KCH strain” of E. adenoideis (see Popstein and Vrba 2011) which possess an oocyst measuring on average 19.4×16.1 μm (SI=1.20). However, the E. adenoideis KCH strain develops within the lower intestinal tract and principally in the ceca.

Discussion
E. meleagrimitis (USMN08-01-Line 5), derived from parasites originally isolated in Minnesota, has been characterized biologically in the present study to complement molecular data published previously (El-Sherry et al. 2013). Tyzzer (1929) first described E. meleagrimitis as a new species. Subsequent descriptions provided additional observations on endogenous development (Hawkins 1952; Clarkson 1959; see Table 5 for summary) and pathogenicity (Clarkson 1959). The dimensions of the oocysts [18.9×15.7 μm (SI=1.21)] agreed with previously reported measurements for E. meleagrimitis: 18×15.25 μm by Tyzzer (1929), 19.17×16.28 μm (SI=1.17) by Hawkins (1952), and 20.1×17.3 μm (SI=1.16) by Clarkson (1959).

According to Clarkson (1959) and Ruff et al. (1980), the first and second asexual generations of E. meleagrimitis were located in principally in upper jejunum of the IT but could be found from the duodenum to just proximal to Meckel’s diverticulum. In contrast, in the present study, first- and second-generation asexual stages were also detected in the lower part of the intestine (ileum, cecal neck, and rectum). In the present study, mature first-generation schizonts were first observed as early as 32 HPI. Clarkson (1959) first detected mature first-generation meronts at 48 HPI (17×13 μm) in the crypts whereas Ruff et al. (1980) found this stage much earlier at 24 HPI. In the study by Ruff et al. (1980), the observed mature first-generation meronts were smaller in size (12×10 μm) and were located within enterocytes lining the sides of villi as opposed to deeply within the crypts as was observed by other workers (Hawkins 1952; Clarkson 1959; present study).

Our observation of gamonts developing as early as 112 HPI following three distinct asexual cycles agrees with the observations of Clarkson (1959), but Ruff et al. (1980) described as many as five asexual generations before the appearance of gamonts. Our observations and those of Clarkson (1959) suggest that perhaps Ruff et al. (1980) were working with a biologically distinct strain or perhaps another species of Eimeria as the latter authors suggested. The monotypic strain of E. meleagrimitis characterized in the present study appears to conform biologically to the original brief description of the parasite by Tyzzer (1929) that was later elaborated upon by Clarkson (1959).

Clarkson (1959) reported the spread of the sexual stages in the whole IT (including the cecum and rectum) with the highest concentration the upper part of the intestine (duodenum and upper jejunum). Interestingly, in the current work, neither gamonts nor any other stages were detected distally within cecal pouches. The relative abundance of sexual stages in the jejunum and rectum compared to the upper and lower duodenum at about 120 HPI agrees with Hawkins’ (1952) observation that the jejunum was parasitized more intensely than other sites and agrees with Tyzzer’s (1929) description of the presence of stages mainly in the
lower part of the small intestine. Detection of oocysts at 120 HPI suggests that the prepatent period for *E. meleagrimitis* USMN08-01-Line 5 is approximately 5 days which agrees with previous descriptions for this species (Clarkson 1959) and our observations on the pattern of oocyst shedding in experimentally infected poults (Fig. 6).
If most enterocytes were infected and killed by early merogonic development, then the later endogenous stages might not find enough healthy enterocytes to complete their life cycle successfully. Paradoxically, an infection started with large numbers of infective oocysts can result in fewer oocysts being shed and less intensive lesions being observed than during infections initiated with fewer infective parasites. Although productive dose thresholds have been estimated for seven *Eimeria* species of chickens (Williams 2001), ours is the first such work for *Eimeria* spp. of turkeys. In the present study, lesions associated with modest infective doses (i.e., $4 \times 10^4$ or $1 \times 10^5$ oocysts/bird) were located mainly in the upper intestinal tract (duodenum and jejunum), the demonstrated predilection site of endogenous development of *E. meleagrimitis*. Increasing infective doses resulted in lesions developing throughout the digestive tract, including the ileum and cecal neck with a decrease in the severity of lesions within the duodenum (Table 4). We interpret this as a response to the exhaustion of suitable enterocytes in the upper intestinal tract (by direct lysis or bystander damage) necessary for continued parasite development; in response, the parasites infect an expanded region of the digestive tract including the ileum and cecal necks (i.e., $2 \times 10^5$ or $5 \times 10^5$ oocysts/bird). Despite high lesion scores and apparently severe mucosal damage, no mortalities due to coccidiosis were observed with doses up to $5 \times 10^5$ oocysts/bird which contrasts sharply with Clarkson’s (1959) observation of up to 36% mortalities in birds infected with only $10^5$ oocysts. Differences in the parasite strain, genetics of the infected turkeys, feed, or age at inoculation, among other factors, may be sufficient to explain Clarkson’s (1959) and our incongruent observations on the pathogenicity of this parasite.

The benchmark species data generated in the present study coupled with previously generated sequence data from two genetic loci [one nuclear-encoded (18S rDNA) and the second mitochondrial-encoded (COI)] from the same isolate (El-Sherry et al. 2013) and deposition of a new name-bearing type will facilitate future studies with the same species and provide a stable “species concept” for *E. meleagrimitis*. Such complementary biological and molecular data will likely be required to resolve the sometimes contradictory literature regarding *Eimeria* spp. infecting turkeys.

**Acknowledgments** This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada; the Ontario Ministry of Agriculture, Food and Rural Affairs; and the Federal Economic Development Agency for Southern Ontario to J.R.B. and scholarship support from the Ministry of Higher Education and Research, Egypt, administered through the Bureau of Cultural and Educational Affairs of Egypt in Canada to S.E.S. Ms. Julie Cobean is thanked for her skilled technical assistance.
References

Long PL, Millard BJ, Shirley MW (1977) Strain variation within Eimeria meleagrimitis from the turkey. Parasitology 75:177–182
Williams RB (2001) Quantification of the crowding effect during infections with the seven Eimeria species of the domesticated fowl: its importance for experimental designs and the production of oocyst stocks. Int J Parasitol 31:1056–1069
APPENDIX 8: THE COMPLETE MITOCHONDRIAL GENOME SEQUENCE OF
Hepatozoon catesbiana (Apicomplexa; Coccidia; Adeleorina), a blood
parasite of the Green frog, Lithobates (Rana) clamitans

(The contents of this chapter have been published as follows: Alexandre N. Leveille, Mosun E. Ogedengbe, Mian A. Hafeez, Hsiang-Hsien(abb) Tu, and John R. Barta (2014). “The complete mitochondrial genome sequence of Hepatozoon catesbiana (Apicomplexa; Coccidia; Adeleorina), a blood parasite of the Green Frog, Lithobates (Rana) clamitans”. Journal of Parasitology 100: 651–656)
THE COMPLETE MITOCHONDRIAL GENOME SEQUENCE OF *HEPATOZOOON CATESBIANAE* (APICOMPLEXA: COCCIDIA: ADELEORINA), A BLOOD PARASITE OF THE GREEN FROG, *LITHOBATES* (FORMERLY *RANA*) CLAMITANS

Alexandre N. Leveille, Mosun E. Ogedengbe, Mian A. Hafeez, Hsiang-Hsien (Abby) Tu, and John R. Barta

Department of Pathobiology, Ontario Veterinary College, University of Guelph, Guelph, ON, N1G 2W1 Canada. Correspondence should be sent to: jbarta@uoguelph.ca

ABSTRACT: A complete mitochondrial genome for the blood parasite *Hepatozoon catesbiana* (Alveolata; Apicomplexa; Coccidia; Adeleorina; Hepatozoidae) was obtained through PCR amplification and direct sequencing of resulting PCR products. The mitochondrial genome of *H. catesbiana* is 6,397 bp in length and contains 3 protein-coding genes (cytochrome c oxidase subunit I [COI]; cytochrome c oxidase subunit III [COIII]; and cytochrome B [CytB]). Sequence similarities to previously published mitochondrial genomes of other apicomplexan parasites permitted annotation of 23 putative rDNA fragments in the mitochondrial genome of *H. catesbiana*, 14 large subunit rDNA fragments, and 9 small subunit rDNA fragments. Sequences corresponding to rDNA fragments RNA5, RNA8, RNA11, and RNA19 of *Plasmodium falciparum* were not identified in the mitochondrial genome sequence of *H. catesbiana*. Although the presence of 3 protein-coding regions and numerous putative rDNA fragments is a feature typical for apicomplexan mitochondrial genomes, the mitochondrial genome of *H. catesbiana* possesses a structure and gene organization that is distinct among the Apicomplexa. This is the first complete mitochondrial genome sequence obtained from any apicomplexan parasite in the suborder Adeleorina.

*Hepatozoon* spp. are heteroxenous apicomplexan parasites (suborder Adeleorina) that use a wide range of vertebrates as intermediate hosts including amphibians, reptiles, birds, and some mammals (Smith, 1996; Baneth et al., 2003). *Hepatozoon* species are characterized by the formation of large, thick-walled poly-sporocystic oocysts found within the tissues of blood-feeding invertebrate definitive hosts such as dipteran insects, ticks, leeches, mites, lice, and fleas that facilitate transmission of the parasite to a vertebrate intermediate host through ingestion of the definitive host (Smith, 1996; Barta, 2000). *Hepatozoon catesbiana* undergoes sexual replication in the gut epithelium of a culicine mosquito (*Culex territans*) definitive host. Consumption of a mosquito containing mature, infective oocysts allows for transmission to the green frog (*Lithobates clamitans*), the intermediate host. Asexual replication then occurs in the hepatocytes of the liver. Resulting merozoites enter erythrocytes and develop into large, crescent-shaped gamonts that remain until they are consumed by blood-feeding mosquitoes. Fully developed gamonts usually displace the nucleus of the host laterally (Desser et al., 1995; Smith, 1996; Barta, 2000). Unfortunately, the presence of these morphologically similar gamonts in infected erythrocytes is common to various hemogregarine sensu lato blood parasites in the genera *Haemogregarina*, *Karyolysus*, *Desseria*, *Cyrilia*, or *Hepatozoon*. Assignment to these genera requires knowledge of the sporogonic development of these hemogregarines within their invertebrate definitive hosts. This requirement precludes assignment of a particular parasite to the appropriate genus by morphological features of blood smear specimens (Barta, 2000), although a blood smear can demonstrate patent infections with a hemogregarine parasite (Desser et al., 1995; Smith, 1996). As well, certain *Hepatozoon* spp. seem to have a low host specificity for either or both their invertebrate and vertebrate hosts. These pleomorphic characteristics make it difficult to properly identify or describe species based solely on morphological features alone (Desser et al., 1995; Smith, 1996).

Mitochondrial genes seem to be useful in phylogenetic comparisons because many are conserved across organisms that use oxidative phosphorylation as a means of respiration. The mitochondrial gene encoding cytochrome c oxidase subunit I (COI) has sufficient sequence variability to effectively delineate recent speciation events and may complement more-slower diverging nuclear genes such as the 18S rDNA (Ogedengbe et al., 2011; Pawlowski et al., 2012). A recent molecular phylogenetic analysis of adeleorinid parasites using 18S rDNA sequences suggested that the genus *Hepatozoon* is more genetically divergent than previously thought and that the genus itself may be paraphyletic (Barta et al., 2012). It is possible that mitochondrial genetic loci may be able to address these open questions by acting as a source of phylogenetically informative sequences such as COI and cytochrome B (CytB).

Formerly, mitochondrial genome sequences have been available for malarial parasites and a few other apicomplexan parasites (e.g., Feagin, 1994; Hikosaka, Watanabe et al., 2011; Feagin et al., 2012). Recently, the genome sequences have become available for piroplasms (i.e., *Babesia* and *Theileria* spp.; Hikosaka et al., 2010) and monoxenous coccidia (various *Eimeria* spp.; Hikosaka, Nakai et al., 2011; Lin et al., 2011), thus providing better insight into the genome evolution, organization, and gene order. Apicomplexan mitochondrial genomes sequenced to date have a wide variability in structure and gene order (Gray et al., 2004; Hikosaka, Watanabe et al., 2011). Apicomplexan mitochondrial genomes have a reduced size (only 6–7 kb) and are known to possess only 3 protein-coding regions: cytochrome c oxidase subunit I (COI); cytochrome c oxidase subunit III (COIII); and cytochrome B (CytB). For parasites in the genus *Hepatozoon*, limited sequence data has been obtained for nuclear ribosomal loci such as 18S rDNA (Barta et al., 2012) or ITS-1 sequences (Boulianne et al., 2007), but no *Hepatozoon* spp. mitochondrial DNA sequences have been published to date. The purpose of this study was to obtain a complete mitochondrial genome sequence for *H. catesbiana*, a blood parasite of the green frog, *Lithobates clamitans*.

**MATERIALS AND METHODS**

**Biological samples and DNA extraction**

Green frogs, *Lithobates clamitans* (Latrielle 1801), were collected by hand from the Speed River, Guelph, Ontario (43°32’51.12”N,
80°11′51.98″W, elevation: 311 m a.s.l.). In the laboratory, blood was collected by puncturing the maxillary facial vein using the methods described by Forzan et al. (2012). A non-heparinized capillary tube was used to collect the blood as it pooled on the surface of the skin. Captured frogs were then returned to their environment. A small portion of each blood sample was smeared on a glass slide. Air-dried thin blood films were stained with a modified Wright's stain and then micrographs of gamonts were taken with a Provis AXT70TRF Digital Photomicroscope (Olympus, Center Valley, Pennsylvania) using bright-field illumination at ×400 magnification (Fig. 1).

DNA was extracted from the collected blood samples using a DNAzo BD kit according to the manufacturer's protocol (BioShop, Molecular Research Center, Inc., Cincinnati, Ohio). After isolation, DNA was quantified spectrophotometrically using a Nanodrop 2000 instrument (Thermo Scientific, Wilmington, Delaware). The resulting DNA sample contained both host and parasite DNA.

**PCR amplification**

Standard PCR was carried out in an MJ Mini thermal cycler (Bio Rad, Hercules, California) in a 50-μl reaction containing 1× Taq buffer, 2 U Platinum® II Taq Polymerase (Invitrogen, Carlsbad, California), 0.8 mM dNTPs, 4 mM MgCl₂, 0.5 μM of each amplification primer (see Table I), and 100–200 ng DNA template (mixed frog-parasite DNA from infected R. clamitans). The PCR reaction profile consisted of an initial melt at 95°C for 10 min followed by 40 amplification cycles (denature at 95°C for 30 sec, anneal at ~50°C [varied depending on primers used] for 15 sec, and extension at 72°C for ~1 min/1 kb of expected amplification product length) and then terminated with a final extension of 72°C for 10 min to complete any partial products. Annealing temperatures were chosen based on Primer3 (Untergasser et al., 2012), implemented from within the primer characteristics provided by Geneious Pro (Version 6.1 and later, available from http://www.geneious.com/).

**Fragment sequencing and analysis**

DNA from the 3 amplified mitochondrial genome fragments was then submitted for sequencing in both directions with the forward and reverse amplification primers using an ABI Prism 7000 Sequence Detection System (Applied Biosystems Inc., Foster City, California) by the Molecular Biology Unit of the Laboratory Services Division, University of Guelph (Guelph, Ontario, Canada). Chromatograms received from sequencing reactions were imported and assembled into contigs using the De Novo Sequence Assembler within Geneious. For the longer fragments, internal sequence primers (Table II) were designed to generate complete sequences of both strands by primer-walking. Ultimately complete, high-quality sequences and collectively spanned the entire mitochondrial genome of *H. catesbianae*.

**TABLE I. PCR amplification primers and resulting fragments used for sequencing.**

<table>
<thead>
<tr>
<th>Fragment</th>
<th>Size (bp)</th>
<th>Primer names</th>
<th>Primer sequences</th>
<th>Anneal temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2,277</td>
<td>Hep_COB_1F</td>
<td>5'-TGGTTTAATGACGCGTCAGGCT-3'</td>
<td>50°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hep_COB-1797R</td>
<td>5'-AGGGTGTCTAGGGTTTACG-3'</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1,466</td>
<td>Hep_C03_182F</td>
<td>5'-GGGGTCGGAATCTTCTCTGTC-3'</td>
<td>50°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hep_COB_164R</td>
<td>5'-GCAATTGGCGACCAAAAAACCA-3'</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4,366</td>
<td>Hep_COB_1020F</td>
<td>5'-GTAGATCAATAACATCCATGT-3'</td>
<td>45°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hep_COB-1038R</td>
<td>5'-CTCCACTATATAAGCGGA-3'</td>
<td></td>
</tr>
</tbody>
</table>

**PCR product purification**

The PCR products were separated electrophoretically using a submersion 0.8% agarose gel with 1× TAE buffer and 5 μl of ethidium bromide dye (10 mg/ml, w/v). An appropriate DNA size standard ladder (Bio Bases, Markham, Ontario, Canada) was used to determine product fragment lengths. Gels were then examined using an ultraviolet transilluminator, and DNA bands of expected sizes were excised using a sterile scalpel. DNA was extracted from the gel slice using the QIAquick Gel Extraction Kit (Qiagen®, Toronto, Canada) according to the manufacturer's instructions.

**Primer design**

All PCR and sequencing primers were developed using Primer3 executed from within Geneious Pro. No degenerate primers for apicomplexan mitochondrial CytB were available, so primers were designed based on an alignment of 36 *Hep* *B* *O* *N* *O* *O* *O* *O* *N* sequences from apicomplexan parasites of various groups (hemopirinids, piromastigotes, tissue coccidia, and enteric coccidia; data not shown). Regions of high sequence identity were used to generate pan-Apicomplexa degenerate primers CytB_400F (5'-CCWDRNGGWCAAATGAGYT-3') and CytB_812R (5'-TAC CAYTCHGHAYDATDG-3'). This primer pair was designed to amplify a 412-nt fragment within the mitochondrial CytB gene from a wide range of Apicomplexa. Sequence data obtained from DNA fragments amplified using the CytB_400F/CytB_812R primer pair were then used to design outward-facing (inverse) primers CytB_H_510F (5'-AAGGCCTG-GATTGGATACGT-3') and CytB_H_540F (5'-CACATTCCG-TAGTGTCCGACTGTA-3'). These primers would be expected to amplify nearly the entire mitochondrial genome of *H. catesbianae*, less the small region (30 bp) between the amplification primers. Using the CytB_H_540F/CytB_H_510R primer pair, an approximately 6.5-kB fragment was amplified weakly, but there was evidence of multiple smaller amplification products as well. Sequencing of the resultant PCR amplification products were found to contain numerous shorter fragments that interfered with successful sequencing by primer-walking along the length of the near-complete mitochondrial genome. Instead, three pairs of *Hepatozoon*-specific PCR primers (Table I) were designed to amplify 3 overlapping fragments (fragments 1, 2, and 3; Fig. 2; Table I) that generated high-quality sequences and collectively spanned the entire mitochondrial genome of *H. catesbianae*. 

**FIGURE 1. Intraerythrocytic gamonts of *Hepatozoon catesbianae* within nucleated erythrocytes of Lithobates clamitans. (A) Mature gamonts measured 23.4 ± 2.32 (18.5–27.1) μm by 5.2 ± 0.52 (4.2–6.2) μm (n = 20) and were found adjacent to an intact host erythrocyte in all cases. (B) Immature gamonts were found in a few frogs, indicating a recent invasion of the erythrocytes by merozoites. Scale bar = 10 μm.
FIGURE 2. Annotated complete mitochondrial genome sequence of *Hepatozoon catesbianae* is illustrated in linear form, but the actual structure of the genome is not known. Three overlapping PCR fragments (fragments 1, 2, and 3) were used to generate the complete genome sequence; overlap regions demonstrated 100% sequence identity between PCR products. Three protein-coding regions (cytochrome *c* oxidase subunit I [COI], cytochrome *c* oxidase subunit III [COIII], and cytochrome B [CyB]) were encoded by the mitochondrial genome of *H. catesbianae*. Numerous putative tRNA fragments were detected in the genome between the protein-coding regions; ribosomal fragment nomenclature follows Feagin et al. (2012).

### Sequence data assembly and annotation of rDNA subunit

Three long open reading frames (ORFs), supposed to represent protein-coding regions, were identified using an ORF search utility within Geneious with a mold/protozoan mitochondrial translation table (trans_t-able 4) to guide the search; these ORFs were translated using the same translation table, and the resulting amino acid (aa) sequences were searched against the public sequence databases using the blastp algorithm to confirm the identities of these 3 protein-coding regions. In addition, transmembrane predictions for the corresponding products from a Hidden Markov Model (HMM) within Geneious and compared to the transmembrane predictions for each translation product were made using TransMem (https://www.geneious.com). Mauve genome alignments were performed in order to estimate the location of the common sequence structures (progressive Mauve algorithm, Darling et al., 2010). All annotated ribosomal subunit fragment sequences were extracted from each of the publicly available annotated genome sequences noted above, and these fragments were aligned pairwise, where possible, to the *H. catesbianae* genome sequence. Any fragments that aligned with the *H. catesbianae* mitochondrial genome with a pairwise sequence identity below 60% were excluding from further analysis. Putative ribosomal fragment annotations follow the nomenclature of Feagin et al. (2012) in their annotation of the mitochondrial genome of *P. falciparum*.

### RESULTS

Intraerythrocytic gamonts of *H. catesbianae* were observed within nucleated erythrocytes of a number of *L. clamitans* with parasitemias that ranged from <1% to ~5%. Mature gamonts were broadly fusiform, lightly basophilic bodies with a wide anterior end that narrowed posteriorly to a short, recurved posterior (Fig. 1A). A central nucleus of diffuse chromatin was evident as well as a number of smaller basophilic granules within the cytoplasm of mature gamonts. Slight sexual dimorphism of gamonts was noted; macrogamonts (Fig. 1A) were slightly more basophilic than were microgamonts, and the nuclei of macrogamonts were more compact and stained more darkly than those of microgamonts. Gamonts measured an average of 23.4 (18.5–27.1) μm by 5.2 (4.2–6.2) μm (n = 20). Gamonts of *H. catesbianae* were found adjacent to a displaced, but intact, host cell nucleus in all cases. In a few frogs, immature gamonts resulting from recent invasion of erythrocytes by merozoites were observed (Fig. 1B). The morphologically similar parasite *Hepatozoon clamatae* (Stebbins 1905) Smith (1996) was observed infrequently infecting frogs at the same locality. Infections with the latter parasite could be distinguished readily by the fragmentation of the nuclei of infected erythrocytes. In a few of the sampled animals, a small number of sporozoites of *Lankesterella minima* (Chaussat 1850) Nöller 1912, an eimeriid coccidium with a heteroxenous life cycle, were observed infecting green frogs concurrently. Blood from frogs infected by more than 1 hemoparasite was not used for generating sequence data from *H. catesbianae*.

The complete mitochondrial genome of *H. catesbianae* had a unit length of 6,397 base pairs (bp) with an overall G+C content of 38.5% (GenBank KF894962). The mitochondrial genome organization consisted of 3 protein-coding genes (COI, COIII, COII).

<table>
<thead>
<tr>
<th>Fragment</th>
<th>Primer name</th>
<th>Primer sequence</th>
<th>Position from genome origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hep_COB-1797R</td>
<td>5'-AGGTGCTCAAGGTCATACCG-3'</td>
<td>219–200</td>
</tr>
<tr>
<td></td>
<td>Hep_COB_1F</td>
<td>5'-TGTTTATACGCAGCTGAGC-3'</td>
<td>4340–4359</td>
</tr>
<tr>
<td></td>
<td>Hep_COB_1020F</td>
<td>5'-GTAGATCAATAACATCCATGT-3'</td>
<td>5333–5353</td>
</tr>
<tr>
<td></td>
<td>Hep_COB_1101R</td>
<td>5'-GTTAGCCAAATCTACGTAC-3'</td>
<td>5413–5395</td>
</tr>
<tr>
<td>2</td>
<td>Hep_COB_182F</td>
<td>5'-GCCGCGGAATCCTCTGTGATC-3'</td>
<td>3037–3056</td>
</tr>
<tr>
<td></td>
<td>Hep_COB_484R</td>
<td>5'-AGGTGCTCAAGGTCATACCG-3'</td>
<td>3855–3833</td>
</tr>
<tr>
<td></td>
<td>Hep_COB_164R</td>
<td>5'-GCCAAATCTACGTCGTTGATT-3'</td>
<td>4502–4482</td>
</tr>
<tr>
<td>3</td>
<td>Hep_COB_1966F</td>
<td>5'-GAGCGGCTTGTAAGAGCACC-3'</td>
<td>6040–6059</td>
</tr>
<tr>
<td></td>
<td>Hep_COB_1799F</td>
<td>5'-ACCTCCCTGTTGATAAGCCGTT-3'</td>
<td>222–243</td>
</tr>
<tr>
<td></td>
<td>Hep_COI_850F</td>
<td>5'-GATATGGCTTGATGGAGTGA-3'</td>
<td>1430–1411</td>
</tr>
<tr>
<td></td>
<td>Hep_COI_400F</td>
<td>5'-GGATATGGACTGGACAGTTGAC-3'</td>
<td>1888–1869</td>
</tr>
<tr>
<td></td>
<td>Hep_COI_1550R</td>
<td>5'-GGGATGGTAGTGGATGATAGT-3'</td>
<td>2749–2728</td>
</tr>
<tr>
<td></td>
<td>Hep_COI_1038R</td>
<td>5'-CTGCAATATATAGACGG-3'</td>
<td>3301–3282</td>
</tr>
<tr>
<td></td>
<td>Hep_COI_1020F</td>
<td>5'-GTTAGATCAATAACATCCATGT-3'</td>
<td>5333–5353</td>
</tr>
</tbody>
</table>

Table II. Sequencing primers.
and CytB), and the G+C content of these protein-coding genes was 37.0% for COI, 35.7% for COIII, and 35.7% for CytB. As in other apicomplexan mitochondrial genomes, many putative rDNA fragments were identified; 14 rDNA fragments of the large ribosomal subunit (LSU) and 9 rDNA fragments of the small subunit (SSU) were identified. Each putative rDNA fragment had sequence similarity to an rDNA fragment that had been functionally annotated in the mitochondrial genome of *P. falciparum* (M76611; see Feagin et al., 2012). Ribosomal DNA fragments encoding RNA5, RNA8, RNA11, and RNA19 in the *P. falciparum* mitochondrial genome were not identified in *H. catesbianae* (Tables III, IV).

The COI coding region of the *H. catesbianae* mitochondrial genome is 1,452 bp in length (bp 2245–794). A possible signal motif CCATTAATTCCA is located just upstream of the coding region that starts with an ATA start codon (transl_table 4 – Mold Protozoan Mitochondrial). The proposed stop codon, TGA, is not typical of the mold protozoan mitochondrial translation table but is used by bacteria (transl_table 11). The COIII coding region is 741 bp in length (bp 2943–3683) and the CytB coding region is 1,113 bp in length (bp 4396–5508). An identical motif (TTTATTGCTATCTATT) is found preceding both the COI and the CytB coding regions and runs directly into a TTA start codon for both sequences. This motif may represent a signal motif (sensu Feagin, 1992). The COIII has a TAA stop codon whereas the CytB concludes at a TAG stop codon (Table III). Interestingly, the first 55 bp of the coding region of both the CytB and COIII reading frames immediately following these possible signal motifs are also identical. This 70-bp duplicated region at the start of the CytB reading frame was in the middle of the 166-bp overlap of fragments 1 and 2. In order to confirm that this was not an experimental error, a 1,730-bp PCR fragment was generated spanning the entire region and subsequently sequenced in both directions (4320–6049 bp; data not shown). Sequencing confirmed that this duplication is present in both the COIII and CytB coding regions. The repeat regions were upstream of the first transmembrane regions predicted by HMM for the *H. catesbianae* CytB and COIII products (see Suppl. Fig. S1).

The translations of COI, CytB, and COIII had 66.7, 62.1, and 47.3% sequence identities with the corresponding translation products of various hemosporidian parasites (a species of *Hepatoctysis*, *Haemoproteus*, and *Plasmodium*, respectively). Transmembrane predictions using HMM showed that the *H. catesbianae* COI translation had the same 12 transmembrane regions predicted for COI of *P. falciparum* (M76611), *E. mitis* (KF501573), and a *Hepatoctysis* sp. (FJ168565) (see Suppl. Fig. S2). Likewise, the CytB translation had 10 transmembrane regions and the COIII translation had 6 transmembrane regions predicted for the corresponding translation products of these same apicomplexan parasites (see Suppl. Figs. S3, S4, respectively).

**DISCUSSION**

Measurements of mature intraerythrocytic gamonts correspond to the dimensions (22.09 ± 1.22 × 5.19 ± 0.52 μm) provided by Desser et al. (1995), and the lack of fragmentation of the host erythrocyte nucleus is in concordance with the species description of *H. catesbianae* (Stebbins 1903) Smith (1996) in ranid hosts (Boulianne et al., 2007).

Amplification of a near-complete mitochondrial genome of *H. catesbianae* using inverse primers proved problematic compared with eimerid coccidia (e.g., *Eimeria* spp. by Lin et al., 2011; Ogedengbe et al., 2013). Generating high-quality sequence
required the amplification of 3 smaller, overlapping fragments of the mitochondrial genome of *H. catesbianae*.

The mitochondrial genome length established in the present study (6,397 bp) is similar to mitochondrial genomes of other apicomplexan parasites (e.g., Feagin et al., 2012). The physical structure of the mitochondrial genome copies of *H. catesbianae* remains unknown, and examining related apicomplexan parasites provides little guidance. Among apicomplexan protists, mitochondrial genome structures are highly diverse, with several structural forms (circular, linear, or linear-concatenated) being reported (Gray et al., 2004; Hikosaka et al., 2010; Hikosaka, Nakai et al., 2011; Hikosaka, Watanabe et al., 2011; Lin et al., 2011; Feagin et al., 2012; Liu et al., 2012; Ogedengbe et al., 2013).

A potential stem-loop structure AGTTAAAAAGAACCATTATACATAGGTTCCTTTATACT (underlined nucleotides potentially involved in base pairing) with no sequence similarity to other sequences from Apicomplexa in GenBank was identified in a non-coding region. For numbering purposes, this structure has been designated as the start of the mitochondrial genome sequence (bp 1–36 of KF894962). Inverted repeat structures such as these in dsDNA can create cruciform hairpin structures and are known to be involved in the initiation of replication and control of gene expression in animal mitochondrial genomes (Clayton, 1982; Sinden, 1994; Boulikas, 1996; Pearson et al., 1996). It is not known whether this structure plays a role in regulation or replication, but this possibility may be worth exploring further.

All sequenced apicomplexan mitochondrial genomes contain 3 protein-coding genes and many rDNA fragments. However, the order and orientation of the 3 protein-coding regions varies considerably among different apicomplexan groups. The genome structure is conserved among closely related parasites such as among piroplasms (Hikosaka et al., 2010), among hemosporinids (Hikosaka, Watanabe et al., 2011), and among eimerid coccidia (Hikosaka et al., 2010; Lin et al., 2011; Ogedengbe et al., 2013). The mitochondrial genome organization of *H. catesbianae* (Fig. 2) was dissimilar to all sequenced mitochondrial genomes in the Apicomplexa (Fig. 3). This structurally distinct mitochondrial genome organization is likely to be shared by other adelie coccidia. To permit comparison among divergent mitochondrial structures, mitochondrial genomes of hemosporinids, eimerid coccidia, and piroplasms were linearized using the start of the COI gene as the first base and were compared to *H. catesbianae* (see Fig. 3).

The mitochondrial genomes of the hemosporinids (*P. falciparum*, *L. sabrazesi*, and *Haemoproteus columbae*) displayed the most overall-arrangement similarities to *H. catesbianae*. The COIII coding region is in the same orientation as the COI CDS; in contrast, the CytB coding region appears to be flipped in both location and orientation. The presence of possible signal motifs upstream of CDS in mitochondrial genomes has not been well documented, although Feagin (1992) noted potential “signal motifs” upstream of the 3 CDS found in the *P. falciparum* mitochondrial genome. Our observation of highly conserved imperfect mirror repeats (sensu Lang, 2005) immediately upstream of putative start codons for both COIII and CytB is suggestive of a transcription regulating sequence (Suplick et al., 1990, Feagin, 1992). Although an identical sequence was not found upstream of the proposed COI start codon, there was an imperfect mirror repeat within the AT-rich region immediately upstream of the start codon for that gene as well (Table III). The COIII/CytB motifs and the COI motif have a consensus sequence of ‘TTTA(T)...(T)ATT’ forming a mirror-repeat structure. The function of such an imperfect mirror repeat, if any, is currently unknown. Similarly, how 2 identical 70-bp regions (2,928–2,997 and 4,381–4,450 of GenBank KF894962) at the start of COIII and CytB CDS arose is unknown. This feature of the *H. catesbianae* mitochondrial genome was also observed in the mitochondrial genome of the closely related adelie parasite *Hepatozoon clamatae* (A. N. Leveille et al., unpub. obs.). The 18 amino acids encoded by the repeat region does not seem to affect the transmembrane structure of the resulting products. Hemosporinids (e.g., *P. falciparum*) have aa sequences that are similarly conserved within the start of their COIII and CytB sequences, especially immediately upstream of the first transmembrane region of each translation product (see Suppl. Fig. S1).

Many of the tentatively identified rDNA fragment locations for *H. catesbianae* agreed with corresponding annotated rDNA fragments of *L. sabrazesi* and *P. falciparum* (see Feagin et al., 2012). The putative LSU and SSU rDNA fragments annotated in the *H. catesbianae* mitochondrial genome were not confirmed functionally (cf. annotations of *P. falciparum* by Feagin et al., 2012). As in all other apicomplexan mitochondrial genomes available, there were no sequences corresponding to tRNAs detected in the mitochondrion of *H. catesbianae*. It is most
probable that tRNAs requisite for mitochondrial conversion are nuclear-encoded (Gray et al., 2004).

The novel structure of the mitochondrial genome of *H. catesbeiana* illustrates the value of exploring less–well-studied apicomplexan parasites. Obtaining primary sequence data and discovering a distinct order and orientation of both protein-coding and putative rRNA-coding regions illustrates that the diversity of mitochondrial genomes available for study within the phylum Apicomplexa has not been exhausted by any means. The mitochondrial genome sequence for this adeleorinid blood parasite can now be used to design appropriate PCR primers to obtain sequence data for COI (or other mitochondrial loci, so-called “DNA barcode” data) from the more than 2,000 adeleorinid parasites that have been described to date. Should the COI locus be found to be a useful genetic target, as was found for the eimerid coccidia (Ogedengbe et al., 2011), DNA barcoding of adeleorinid parasites could be used to help define host specificity and perhaps aid in identifying invertebrate definitive hosts involved in the complicated life cycles of these heteroxenous parasites (Smith, 1996). Ultimately, resolving the evolutionary relationships among adeleorinid parasites, and between these parasites and other Apicomplexa or related alveolate taxa (both parasitic and free-living protists), may require a species-level molecular marker such as the mitochondrial COI locus as well as a more-slowly evolving genetic marker such as nu 18S rDNA (Ogedengbe et al., 2011; Pawlowski et al., 2012).

ACKNOWLEDGMENTS

This research was supported by a Discovery Grant to J.R.B. from the Natural Sciences and Engineering Research Council of Canada. A.N.L. is supported by an OVC Scholarship (MSc) from the Ontario Veterinary College, University of Guelph. Dr. Shiem El-Sherry is thanked for her help assisting with the fieldwork.

LITERATURE CITED


Re-description of a genetically typed, single oocyst line of the turkey coccidium, *Eimeria adenoeides* Moore and Brown, 1951

S. El-Sherry · M. E. Ogedengbe · M. A. Hafeez · M. Sayf-Al-Din · N. Gad · J. R. Barta

Received: 19 March 2014 / Accepted: 31 July 2014 / Published online: 17 August 2014
© Springer-Verlag Berlin Heidelberg 2014

**Abstract** The Guelph strain of *Eimeria adenoeides* was obtained from a commercial turkey flock in Ontario, Canada, in 1985. Single oocyst derived lines of *E. adenoeides* were propagated, and one of them used to re-describe biological and morphological features of *E. adenoeides* in the turkey. Oocysts of this strain are within the lower size ranges in the original species description reported by Moore and Brown (1951); oocysts of the Guelph strain averaged $18.7 \pm 1.4 \, \mu m$ (16.7–22.5) by $14.3 \pm 0.9 \, \mu m$ (13–16.2, *n* = 30) with a shape index (SI) of 1.3±0.1. It is possible that the original species description was based, at least in part, on a mixed culture of two or more *Eimeria* species. Immature first-generation meronts of *E. adenoeides* Guelph strain were observed histologically at 32 h post-infection in the ileum and cecal neck. Early studies reported only two asexual generations suggested that first asexual cycle observed at 32 h post-infection was overlooked. In the present study, three asexual generations were observed before the start of gametogony. The Guelph strain is also characterized by a pre-patent period of 112 h. The Guelph strain of *E. adenoeides* is a highly pathogenic coccidium that forms classic cecal lesions, including prominent caseous cecal cores, during moderate to severe infections. The maximum output of oocysts ($1.77 \times 10^7$ per bird) was obtained from birds inoculated with $1 \times 10^3$ oocysts; maximum fecundity ($1.55 \times 10^5$ oocysts shed per oocyst inoculated) was obtained with an inoculation of $1 \times 10^2$ oocysts, but fecundity dropped dramatically as the inoculation dose increased. To promote stability of the *E. adenoeides* species concept, neotype specimens (a parahapantotype slides series and phototype) have been designated and deposited for future reference.

**Keywords** Coccidiosis · Life cycle · Endogenous development · Fecundity · 18S rDNA · Mitochondrial cytochrome *c* oxidase I · Taxonomy

**Introduction**

Seven species from genus *Eimeria* are responsible for inducing coccidial infections in turkeys. Each invades a specific location in the intestine of affected birds. *Eimeria meleagrimitis*, *Eimeria dispersa*, *Eimeria innocua*, and *Eimeria subrotunda* invade the upper part of the turkey gastrointestinal tract (GIT) (e.g., duodenum, jejunum, and upper ileum) primarily, whereas *Eimeria meleagridis*, *Eimeria gallopavonis*, and *Eimeria adenoeides* affect mainly the lower part of the digestive tract (e.g., lower ileum, cecum, and rectum) (Clarkson and Gentles 1958; Chapman 2008). *E. adenoeides* is considered the most pathogenic *Eimeria* species affecting the lower part of the turkey intestine (Lund and Farr 1965). Moore and Brown (1951) isolated and formally described *E. adenoeides* from a field sample isolated in turkey farm in New York, USA. They provided only limited information about the morphometrics of oocysts, pre-patent period, gross and microscopic lesions, and a short description of the life cycle stages and their distributions. Clarkson (1956, 1958) made considerably more detailed studies on an *E. adenoeides* strain that had been identified, isolated, and purified in England including detailed observations on the life cycle stages and pathogenicity of the latter isolate. Both isolates were reported to possess ellipsoidal oocysts that tend to be slightly narrower at one end with the presence of a refractile globule at that smaller end with a wide range of...
dimensions reported (18.95–31.26 by 12.60–20.94 μm according to Moore and Brown 1951). Clarkson (1958) described the presence of at least two asexual stages. The first generation observed was located in the cecal neck and ileum. Clarkson did not detect any mature schizonts until 60 h post-infection (HPI), and these were found within enterocytes lining the sides of the villi and not in enterocytes lining the crypts. Parasites in the second asexual generation and gametogenic stages were reported to infect both villar and crypt enterocytes (Clarkson 1958). According to Moore and Brown (1951) and Clarkson (1956), *E. adenoeides* infections were confined to the lower GIT (lower ileum, cecum, and rectum). Clinical signs were reported to include depression, orange mucoid watery diarrhea with small quantities of blood, and dehydration. The lower part of the intestine, especially the two ceca, become congested and swollen with petechial hemorrhage and, ultimately, formation of a yellow cheesy (caseous) cecal core at 6 to 7 days post-infection.

The severe pathogenicity of *E. adenoeides* could be expected to have an economic impact on turkey production. Clarkson (1958) found significant reduction in weight gain and less efficient feed conversion in poults infected with a moderate dose of oocysts (2.5×10^4). A higher dose (2×10^5) resulted in 100 % mortality in the same age of poults. Heins (1969) observed 100 % mortalities of poults infected as few as 8×10^4 oocysts/bird, and even low infectious doses (5×10^3 oocysts/bird) reduced weight gains significantly. As Chapman (2008) noted, birds of different ages or breeds on different feed rations may respond differently to the same coccidial challenge, so it can be difficult to make comparisons between studies.

Like all coccidia, *E. adenoeides* was described using summation of morphometric and biological information. The stability of these diagnostic features was questioned recently when Popstein and Vrba (2011) isolated two strains of *E. adenoeides* that differed significantly in oocyst morphology. They described the *E. adenoeides* KR strain as possessing large ellipsoidal oocysts measuring ∼27.8 × 19.4 μm and a second *E. adenoeides*, the KCH strain, possessing a smaller oval oocyst ∼19.4 × 16.1 μm. This was the first time that such wide morphometric variation (43 % variability in the length) had been reported between two strains of a single species. They reported similar pathogenicity in both strains and appreciable immunological cross-protection in turkey pouls between them (77–95 % reduction in oocyst output following heterologous challenge in immunized birds). These observations bring into question the efficacy of the traditional ways of identifying *Eimeria* species, specifically oocyst morphometrics and species-specific immunity in the natural host (Chapman 2008).

The present work characterizes a single-oocyst-derived strain of *E. adenoeides* that was obtained from a commercial turkey flock in Ontario, Canada, in 1985. Oocyst morphometrics, endogenous and exogenous biological features of this strain, as well as oocyst shedding patterns and overall parasite fecundity are presented to complement previously reported molecular characterization of the *E. adenoeides* Guelph strain (El-Sherry et al. 2013). The present study addresses the lack of stability of the species concept for *E. adenoeides* Moore and Brown (1951) by providing a formal taxonomic re-description for this important pathogen of turkeys, including designation and deposition of neotypes linked to multilocus genetic markers.

### Materials and methods

#### Origin and derivation of single-oocyst-derived lines of *E. adenoeides*

A strain of *E. adenoeides* was obtained from a commercial turkey flock in Ontario, Canada, in 1985. Single-oocyst-derived lines of *E. adenoeides* were propagated in specific parasite-free poults in CAF Isolation unit of the Campus Animal Facility, University of Guelph (Guelph ON, Canada). Turkey pouls were provided feed and water ad libitum; all experimental manipulations were reviewed and approved by the University of Guelph’s Animal Care Committee and complied with the Canadian Council on Animal Care’s Guide to the Care and Use of Experimental Animals (second edition). Single-oocyst lines (lines 3, 4, 5, and 7) were derived from the *E. adenoeides* isolate by following the method of Remmler and McGregor (1964) with the modification of using agar pieces to carry a single oocyst for delivery to the bird orally within gelatin capsules. Line 5 was used for all experiments described herein. The sequence of the mitochondrial cytochrome *c* oxidase subunit *I* (mtCOI) and nuclear small subunit 18S rDNA (nu 18S) loci of this single-oocyst-derived line of *E. adenoeides* has been previously reported by El-Sherry et al. (2013). Sporulation of the oocysts were conducted by suspending them in 2.5 % potassium dichromate and incubated at 26 °C on a rotary shaker for a minimum of 72 h. Oocysts were then washed from the potassium dichromate by centrifugation (1,000× *g* for 10 min), re-suspended in 1× phosphate buffered saline (PBS), and stored at 4 °C. Before experimentation, oocysts were collected by centrifugation (2,500× *g* for 10 min) and re-suspended in distilled water. Sporulated oocysts were counted and diluted to the desired doses (Shirley 1995). Oocysts were stored for less than a month before in vivo experimental use.

#### Morphometrics of the oocyst

A Provis AX70 photomicroscope (Olympus Canada, Richmond Hill, ON, Canada) fitted with a digital imaging device (Infinity3-1C, Lumenera Corporation, Ottawa, ON, Canada) was used for all experiments described herein. The sequence of the mitochondrial cytochrome *c* oxidase subunit *I* (mtCOI) and nuclear small subunit 18S rDNA (nu 18S) loci of this single-oocyst-derived line of *E. adenoeides* has been previously reported by El-Sherry et al. (2013). Sporulation of the oocysts were conducted by suspending them in 2.5 % potassium dichromate and incubated at 26 °C on a rotary shaker for a minimum of 72 h. Oocysts were then washed from the potassium dichromate by centrifugation (1,000× *g* for 10 min), re-suspended in 1× phosphate buffered saline (PBS), and stored at 4 °C. Before experimentation, oocysts were collected by centrifugation (2,500× *g* for 10 min) and re-suspended in distilled water. Sporulated oocysts were counted and diluted to the desired doses (Shirley 1995). Oocysts were stored for less than a month before in vivo experimental use.
Canada) controlled using iSolution Lite image analysis software (Hoskin Scientific, Burlington, ON, Canada) were used to measure the lengths and widths of 30 oocysts and 30 sporocysts of *E. adenoeides*. Measurements of each life cycle stage (*n* = 30/stage) were similarly obtained. All measurements are reported as means ± standard deviation followed by range in parentheses.

**Characterization of the endogenous life cycle stages**

Twelve 18-day-old poults (Hybrid Turkeys, Kitchener, ON, Canada) were infected individually with increasing oocyst doses every 8 h and then killed humanely at the same time to obtain tissue samples throughout the endogenous development of this strain (see Table 1). For each bird, the intestinal tract was removed, and 1-cm-long pieces were taken from eight locations: (1) middle of the descending duodenum; (2) middle of the ascending duodenum; (3) jejunum approximately 3 cm proximal to Meckel’s diverticulum; (4) jejunum approximately 3 cm distal to Meckel’s diverticulum; (5) mid-point of the ileum; (6) proximal 1 cm of the cecal neck; (7) middle of the cecal pouch; and (8) middle of the rectum. Samples were rinsed briefly with saline and then placed in freshly prepared Serra fixative solution (absolute ethanol [60 %/v/v], 37 % formaldehyde [30 %/v/v], and glacial acetic acid [10 %/v/v]) for 24 h at room temperature. Samples were then embedded in paraffin, sectioned at a thickness of ∼4 μm, mounted on a glass slide, and then stained with hematoxylin and eosin (Animal Health Laboratory, University of Guelph). Fixation and processing of all tissue samples from all birds proceeded in parallel to minimize variation among samples. The resulting histological sections were examined for the presence of endogenous parasite stages.

**Oocyst inoculum required to achieve maximal oocyst production**

Thirty 3-week-old coccidia-free turkey poults were sorted into six groups with five birds in each group. Each group of five poults was placed into a cage with raised wire flooring, and aluminum foils were placed beneath the cage floor for daily fecal collection. Birds in each group were given doses of sporulated oocysts in 1 ml PBS via oral gavage as follows—1 × 10^2, 1 × 10^3, 5 × 10^3, 1 × 10^4, 2.5 × 10^4, or 5 × 10^4 oocysts. Total daily output of oocysts was determined from days 4–9, inclusive. The total number of oocysts found within each sample covering a 24-h collection period was determined using a McMaster counting technique with saturated NaCl as the flotation medium (Long et al. 1977), and total oocyst output for each group of birds receiving a particular inoculation dose was determined by totaling the counts for all days. The number of oocysts shed per bird was calculated by dividing the total oocysts shed by the number of birds contributing to those samples. All birds used to determine oocyst shedding and fecundity were killed humanely at the conclusion of fecal collection to ensure that all cecal core material had been released from each bird. Fecundity was calculated by dividing the calculated total oocyst production per bird by the dose of oocysts administered.

**Results**

**Species re-description**

**Taxonomic summary:**

 Apicomplexa
  Conoidasida
  Coccidia
  Eucoccidiorida
  Eimeriorina
  Eimeriidae
  *Eimeria adenoeides* Moore and Brown 1951

Geographic origin and date of isolation of parent culture: Guelph, Ontario, Canada, 1985.

Type-host: *Meleagris gallopavo* (Aves, Galliformes, Phasianidae, Meleagridinae).

Type locality: unknown turkey farm in New York State, USA, near 40.71435, −74.00597; see Moore and Brown 1951

Other localities: Guelph, Ontario, Canada; likely cosmopolitan in turkey flocks

**Description of sporulated oocysts**

Oocysts are oval in shape with a doubled-contoured oocyst wall (Fig. 1a), the sporulated oocysts 18.7 ± 1.4 μm (16.7–22.5) by 14.3 ± 0.9 μm (13.0–16.2, *n* = 30) with a SI of 1.3 ± 0.1; and the single refractile polar granule in sporulated oocysts (Fig. 1b). Sporocysts measure 9.7 ± 0.7 μm (9.4–11.2) by 5.9 ± 0.5 μm (4.8–6.7, *n* = 30). The scatter plot of lengths and widths for the measured oocysts demonstrated the presence of a morphometrically homogeneous population (Fig. 2).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Inoculating doses and timing of experimental infections with <em>E. adenoeides</em> (Guelph Strain) in 18-day-old poults used to examine endogenous development of the parasite in the turkey, <em>Meleagris gallopavo</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bird number</td>
<td>Inoculation dose (oocysts/bird)</td>
</tr>
<tr>
<td>1, 2, 3, 4</td>
<td>2 × 10^6</td>
</tr>
<tr>
<td>5, 6, 7</td>
<td>2.6 × 10^5</td>
</tr>
<tr>
<td>8, 9, 10</td>
<td>2 × 10^5</td>
</tr>
<tr>
<td>11, 12</td>
<td>2.5 × 10^4</td>
</tr>
</tbody>
</table>
Endogenous development

The life cycle stages of *E. adenoeides* that were observed microscopically and their locations throughout endogenous development were summarized in Table 2 and illustrated diagrammatically (Fig. 3). No parasites were observed in the duodenum, jejunum, or upper ileum at any time during experimental infections.

**First asexual stage** At 32 h post-inoculation, sporozoites, trophozoites, and immature meronts were located principally within enterocytes lining the sides of the villi, mainly in ileum and cecal neck (Fig. 4a), although a few immature meronts were detected in the rectum. However, none of them were detected in the enterocytes of the middle and distal regions of the ceca (i.e., deeper than the cecal neck). The growth of the first-generation meronts below the host cell nuclei did not appear to alter the position of the nucleus. The earliest apparently maturing first-generation meronts were detected at 32 HPI, and each contained ~45–50 merozoites. Mature meronts measured 12.9±1.9 μm (10.6–16.1) by 10.3±1.6 μm (9.1–14.7). Maturing first-generation meronts continued to be observed until 48 HPI.

**Second asexual stage** Second-generation meronts were first detected at 48 HPI, mainly in the mid-cecum in enterocytes along the sides of the villi but, unlike first-generation meronts, the second-generation meronts were also observed infecting crypt enterocytes (Fig. 4b). Meronts developed near the tips of epithelial cells above the host cell nucleus. Few were detected in the ileum or cecal neck. Mature second-generation meronts were numerous at 56 HPI, and each measured, on average, 9.1±1 μm (8.1–11.3) by 7.3±0.5 (6.4–8.3) μm and contained ~12–16 merozoites. No mucosal damage was evident during the growth and maturation of the second-generation meronts even in regions where they were numerous. Mature second-generation meronts continued to be observed until 64 HPI.

**Third asexual stage** Trophozoites of the third generation were detected at 64 HPI infecting enterocytes along the sides and tips of the villi, as well as within cyst enterocytes of the cecal neck and in the cecum. Mature third-generation meronts detected at 80 and 88 HPI measured on average 8.3±0.8 μm (7.4–9.8) by 7±0.8 μm (6.1–8.9) (Fig. 4c). Each mature third-generation meront contained ~6–8 merozoites. This stage was detected until 88 HPI. Unlike second-generation meronts, there was considerable damage to the epithelium during third-generation development in the cecal neck and cecum.

**Gametogony** Immature gamont stages were first observed at 96 HPI in the cecal neck, cecum, and rectum. Gamonts were observed in enterocytes lining the entire mucosal surface from the crypts to the villar tips (Fig. 4d). Apparently, fully differentiated male and female gamonts were detected as early as 104 HPI, but no unsporulated oocysts were detected within the tissues at this time. By 112 HPI, mature gamonts were common, and oocysts were first detected within the epithelium and lumen of the cecum and rectum. Mature macrogamonts measured, on average, 13.2±1.2 μm (11.5–15.2) by 10.9±0.9 μm (8.9–12.7). Mature microgamonts measured, on average, 11.7±1.1 μm (9.5–13.6) by 9.9±1.5 μm (8.2–13.4). Both mature gamonts and oocysts were no longer observed at 120 HPI.

**Reproductive potential and dose-dependent fecundity**

The relationship between various inoculation doses and the resultant number of oocysts shed by a single 3-week-old poult is illustrated (Fig. 5a). Maximum oocyst output of ~1.7×10⁷ oocysts/bird was obtained from birds inoculated with 1×10⁴ oocysts per os; birds inoculated with higher (5×10³ oocysts/bird) or lower (1×10² oocysts/bird) numbers of oocysts per os shed fewer total oocysts. The total oocyst yield was even
lower with larger inoculum doses (10^4 to 5×10^4 oocysts). Fecundity decreased dramatically with increasing numbers of oocysts in the inoculum (Fig. 5b) from a peak of 1.55×10^5 oocysts generated by each inoculated oocyst at a dose of 1×10^2 oocysts to a low of 1.63×10^2 oocysts generated from each inoculated oocyst when the inoculating dose was 5×10^4 oocysts.

Pathological lesions

This is a highly pathogenic coccidium that can cause necrosis, sloughing of the epithelium, edema, a characteristic white caseous core in the ceca (consists of dead cells as well as gamonts and oocysts), and limited petechial hemorrhage but without the severe blood loss that is usually observed in cecal coccidiosis in the chicken (Fig. 6a, b).

Reference molecular data

Mitochondrial cytochrome c oxidase subunit I (mtCOI) A portion of the mtCOI gene from the single-oocyst derived line of *E. adenoeides* used in the present re-description was amplified using PCR, gel-purified, and sequenced directly (KC346359) as described by El-Sherry et al. (2013); similarly, the partial mtCOI gene from three additional single-oocyst-derived lines of *E. adenoeides* isolated from the same parent line of *E. adenoeides* Guelph strain were likewise amplified by PCR and directly sequenced by the latter authors. All four partial mtCOI sequences were identical to one another (KC346357 to KC34360), as well as to the published partial mtCOI gene sequence of *E. adenoeides* strain KCH (FR846202, Poplstein and Vrba 2011). Mitochondrial COI is the preferred molecular marker for identification of this species.

Nuclear 18S (small subunit) rDNA (nu 18S) Nuclear 18S rDNA sequences could not be obtained by PCR amplification followed by direct sequencing of the PCR product (see El-Sherry et al. 2013). Consequently, nuclear 18S rDNA
sequences (9 near-complete 18S rDNA sequences, KC305177 to KC305185) from the single-oocyst derived line of *E. adenoeides* used in the present biological re-description were obtained by PCR amplification followed by gel-purification, cloning, and sequencing of individual cloned PCR products (see El-Sherry et al. 2013). These authors also sequenced an additional eight near-complete 18S rDNA clones of the PCR product amplified from a second single-oocyst-derived line of *E. adenoeides* (KC305169 to KC305176) that was isolated from the same parent line of *E. adenoeides* Guelph strain. All sequences should, collectively, be considered valid reference sequences for *E. adenoeides* Guelph strain (see El-Sherry et al. 2013).

**Designation of a neotype and type deposition**

The neotype consists of a hapantotype series of four histological preparations of the intestinal tract of experimentally infected turkeys and a phototype of sporulated oocysts as follows: CMNPA 2014-0004.1, 48 HPI demonstrating mature first-generation meronts; CMNPA 2014-0004.2, 56 HPI demonstrating mature second-generation meronts; CMNPA 2014-0004.3, 96 HPI demonstrating mature third-generation meronts; CMNPA 2014-0004.4, 112 HPI demonstrating mature macrogamonts, microgamonts, and unsporulated oocysts; and CMNPA 2014–0004.5, a phototype of the sporulated oocyst of this parasite.

**Discussion**

Differentiation among the three *Eimeria* species parasitizing the lower part of the intestinal tract in turkeys by dependence on only oocyst morphometrics and shedding patterns has proven to be difficult. These *Eimeria* spp. are reported to have similar average oocyst dimensions (i.e., 24.4×18.1 μm, SI 1.35, for *E. meleagridis*, Tyzzer 1929; 25.6×16.6 μm, SI 1.54 for *E. adenoeides* according to Moore and Brown in 1951; and 27.1×17.2 μm, SI 1.52, for *E. gallopavonis* according to Hawkins in 1952) that has made it challenging to distinguish among these parasites based solely on oocyst dimensions. For example, in the species description of *E. adenoeides* by Moore and Brown (1951), the authors recorded a wide range of dimensions for their new species (18.9 to 31.3 μm in length; 12.6 to 20.9 in width). Isolation and purification of five different single oocysts lines from five *Eimeria* spp. infecting turkeys produced oocysts that are highly conserved in shape and dimensions within each line (El-Sherry et al. 2014). These results suggest that pleomorphism in oocyst dimensions...
within a single species is unlikely to explain the wide ranges of dimensions previously recorded for some *Eimeria* spp. infecting turkeys, notably *E. adenoeides*. The more likely explanation is that the original species description of *E. adenoeides* by Moore and Brown (1951) was complicated by mixed cultures of two or more *Eimeria*. This possibility may be observed in Fig. 1 of Moore and Brown (1951) in which oocysts with divergent dimensions and morphological features were claimed to represent a single oocyst line of a single species. Even in the Clarkson’s later (1958) characterization of a strain of *E. adenoeides* that was isolated in England, an apparently mixed culture was illustrated (see Fig. 5 of Clarkson 1958), and the histogram of oocyst dimensions showed at least two apparent populations of oocysts (see Fig. 5 of Clarkson 1958). Clarkson (1958) also reported two different prepatent periods (a few birds started shedding oocysts at about 114 h while most of the birds began shedding at 132 h) when turkeys were infected with the “single species” line of *E. adenoeides*.

The smaller dimensions reported in the species description by Moore and Brown (1951) largely agree with the dimensions of the *E. adenoeides* Guelph strain (19±1.4 μm (17–22) by 14±0.9 μm (13–16, n=30) with SI of 1.3±0.1), similar to the KCH strain of Poplstein and Vrba (2011); the larger end of the dimensions reported in the original species description agree with the dimensions of *E. gallopavonis*. In a recent study (El-Sherry et al. 2014) using sequence-based genotyping using mtCOI, *E. adenoeides* Guelph strain had a COI sequence that was identical to the *E. adenoeides* KCH strain, and the mtCOI sequence of *E. meleagridis* USAR97-01 matched that of *E. adenoeides* KR strain isolated by Poplstein and Vrba (2011).

Although the mean oocyst dimensions of the Guelph strain of *E. adenoeides* were at the lower end of the size ranges found in the original species description, the life cycle stages of *E. adenoeides* Guelph strain and their distributions in the digestive tract of infected turkeys largely agree with the limited information on endogenous development in the original species description of *E. adenoeides* by Moore and Brown (1951). The endogenous development of the Guelph strain also matched observations by Clarkson (1958) in his more detailed study. The original species description by Moore and Brown (1951) did not provide any information about the life cycle stages before the fifth day post-infection. Both earlier studies reported endogenous development confined to the lower ileum, ceca, and rectum of infected birds.

![Fig. 4](image_url) **Endogenous development of *E. adenoeides* in turkey poults:** a 40 HPI, large first-generation meronts (arrow) located within enterocytes lining the sides of villi mainly in the ileum and cecal necks. First-generation stages developed below the host cell nuclei but did not appear to alter the cell shape; b 56 HPI, mature second-generation meronts (arrows) mainly in the mid-cecum within enterocytes lining the sides of villi but also within crypt enterocytes. Stages developed near the mucosal surface of epithelial cells above the host cell nucleus; c 66 HPI, mature third-generation (arrows) observed in the crypts as well as along the sides of the villi and villar tips of the cecal neck and in the cecum; d unsporulated oocyst (oo), developing microgamonts (mi), and mature macrogamont (ma), were observed at 112 HPI. Scale bar=20 μm
According to Clarkson (1958), the first generation was observed in the cecal neck and ileum alongside the villi but not within the middle and distal regions of the ceca (i.e., deeper than the cecal neck). Although mature first-generation stages were not observed until 60 HPI, he described the schizonts as being large (30×18 μm) and containing up to 700 merozoites. In the present study, the first-generation immature meronts were detected as early as 32 HPI alongside the villi mainly in ileum and cecal neck. First-generation merogonic stages were not observed infecting enterocytes of the middle and distal regions of the ceca. Mature meronts within villar enterocytes measured 13×10 μm and consisted of only 45–50 merozoites. Clarkson (1958) reported the presence of at least two asexual small stages before the start of gametogony. Three asexual generations were observed before the start of gametogony with an average duration of each asexual cycle of 16 h.

A summary of descriptions of the life cycle stages in the original species description (Moore and Brown 1951) and a subsequent detailed biological study (Clarkson 1958) compared with observations in the present study is found in Table 3. Interestingly, the times post infection (HPI) for the development of the two asexual stages observed by Clarkson (1958) was similar to what was observed for the second and third asexual generations for the Guelph strain. According to Clarkson (1958), stages of the first asexual generation were not detected before 60 HPI, but this was nearly the same time that the second asexual generation was observed in present
study. Likewise, the stage that was described to be the second generation in Clarkson's (1958) work was observed at about 96 HPI which was nearly the same time when the development of the third generation observed in the present study. It is possible that Clarkson (1958) did not observe the presence of the first generation which we detected as early as 32 HPI, and the stages that were described by him as first and second generations were actually the second and the third generations. Clarkson (1958) injected merozoites from these “second-generation” meronts into coccidia-free birds and observed that these merozoites developed directly into sexual stages; he consequently concluded that there are only two asexual developmental stages for this *Eimeria* species. However, the results of the present study suggested that Clarkson (1958) did not detect the actual first-generation asexual stages and interpreted the second-generation stages to be first-generation stages in his report.

*E. adenoeides* has been described as one of the most pathogenic species that can infect the turkey (Chapman 2008). The pathological lesions that were observed for the *E. adenoeides* Guelph strain are considered pathognomonic for cecal coccidiosis caused by *E. adenoeides* (e.g., Chapman 2008). The presence of hemorrhage and caseous exudate was observed less commonly in the ileum or rectum, even though histological examination showed that these areas were heavily parasitized as well. In the present study, hemorrhage due to *E. adenoeides* in the ceca of turkeys was much less noticeable than during infections by *Eimeria tenella* in the ceca of chickens. Unlike *Eimeria* spp. in chickens, the number of parasites that invade the intestinal tissue does not correlate well with the severity of associated pathological lesions; this inconsistency seems to be a common phenomenon for *Eimeria* spp. that infect turkeys (e.g., Ruff et al. 1976; Bemrick and Hammer 1978). For example, Madden and Ruff (1979) compared the appearance of the intestinal mucosa
Table 3  Comparison of endogenous development of *E. adenoeides* as described by various authors

<table>
<thead>
<tr>
<th>Features</th>
<th>Previous and present observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Moore and Brown 1951</td>
</tr>
<tr>
<td></td>
<td>Clarkson 1958</td>
</tr>
<tr>
<td></td>
<td>Present study</td>
</tr>
<tr>
<td>Oocyst size (L × W)</td>
<td>25.6 μm (18.95–31.26) by 16.59 μm (12.60–20.94)</td>
</tr>
<tr>
<td></td>
<td>25.6×16.25 μm</td>
</tr>
<tr>
<td></td>
<td>18.7±1.4 μm (16.7–22.5) by 14.3±0.9 μm (13–16.2)</td>
</tr>
<tr>
<td>Prepatent period</td>
<td>112 h</td>
</tr>
<tr>
<td></td>
<td>132 h</td>
</tr>
<tr>
<td></td>
<td>112 h</td>
</tr>
<tr>
<td>Mature first-generation</td>
<td>Size</td>
</tr>
<tr>
<td></td>
<td>Generally, stages are present in lower third of small intestine, ceca and rectum. Within villar and crypt enterocytes</td>
</tr>
<tr>
<td></td>
<td>Mainly in ileum and cecal neck, fewer in rectum</td>
</tr>
<tr>
<td></td>
<td>Below the host cell nuclei along the sides of the villi</td>
</tr>
<tr>
<td></td>
<td>Intestinal location</td>
</tr>
<tr>
<td>Number of merozoites</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>~45–50</td>
</tr>
<tr>
<td>32 HPI</td>
<td></td>
</tr>
<tr>
<td>Mature second-generation</td>
<td>First observed</td>
</tr>
<tr>
<td></td>
<td>~700 merozoites</td>
</tr>
<tr>
<td></td>
<td>12–16</td>
</tr>
<tr>
<td></td>
<td>60 HPI</td>
</tr>
<tr>
<td></td>
<td>56 HPI</td>
</tr>
<tr>
<td>Mature third-generation</td>
<td>First observed</td>
</tr>
<tr>
<td></td>
<td>~12–24</td>
</tr>
<tr>
<td></td>
<td>6–8</td>
</tr>
<tr>
<td></td>
<td>96 HPI</td>
</tr>
<tr>
<td></td>
<td>80 HPI</td>
</tr>
<tr>
<td>Mature gamonts</td>
<td>First observed</td>
</tr>
<tr>
<td></td>
<td>Male and female 20×18 μm</td>
</tr>
<tr>
<td></td>
<td>Macrogamonts 13.2×10.9 μm; microgamonts 11.7×9.9 μm</td>
</tr>
<tr>
<td></td>
<td>Posterior third of small intestine, ceca, and rectum</td>
</tr>
<tr>
<td></td>
<td>Cecal neck, middle and distal ceca, and rectum</td>
</tr>
<tr>
<td></td>
<td>Along the sides of the villi up to the tips as well as deeply within the crypts</td>
</tr>
<tr>
<td></td>
<td>Intestinal location</td>
</tr>
<tr>
<td></td>
<td>Within villi, villar tips, and crypts</td>
</tr>
<tr>
<td></td>
<td>Along the sides of the villi up to the tips as well as deeply within the crypts</td>
</tr>
<tr>
<td></td>
<td>Histological location</td>
</tr>
<tr>
<td></td>
<td>In the crypts as well as along the sides the villi and in the crypts</td>
</tr>
<tr>
<td></td>
<td>In the crypts as well as along the sides the villi and in the crypts</td>
</tr>
<tr>
<td></td>
<td>First observed</td>
</tr>
<tr>
<td></td>
<td>114 HPI</td>
</tr>
<tr>
<td></td>
<td>104 HPI</td>
</tr>
</tbody>
</table>
noted that increasing the inoculating doses to $2 \times 10^5$ sporulated oocysts/bird still failed to produce extensive mucosal disruption in the cecal neck and rectum.

Two other *Eimeria* species that infect the lower intestinal tract of turkeys can cause similar macroscopic lesions, although not usually associated with the large cecal core illustrated in Fig. 6. The pathological features of *E. gallopavonis* infections can be difficult to differentiate from *E. adenoeides*. According to Lund and Farr (1965), *E. gallopavonis* should be considered to be a highly pathogenic *Eimeria* sp. of turkeys as well. Edema, ulceration, inflammation, and caseation of the ileum and ileo-cecal junction can be associated with *E. gallopavonis* infections (Reid 1972; Joyner 1973). In heavy infections, caseous material may accumulate in the cecal pouches as well, but intracellular development of this parasite within the ceca may be restricted to the cecal neck only. For *E. meleagridis*, Clarkson (1959) reported the infrequent presence of caseous material and plugs in feces and within the cecal pouch, but he concluded that this species is comparatively non-pathogenic. A more recent study by Matsler and Chapman (2006) concurred and described *E. meleagridis* as mildly pathogenic due to its modest negative effects on weight gain and feed conversion ratio in infected pouls.

Live coccidiosis vaccines marketed for turkeys will likely include *E. adenoeides* because of its demonstrated pathogenicity to pouls (present study) and its ability to readily elicit a protective immune response to homologous challenge infections (data not shown). Information regarding the maximum production dose and fecundity of this species (Williams 2001) is essential for use of this parasite in a live, virulent vaccine so that vaccination is achieved without causing clinical effects even following oocyst “cycling” in litter (Chapman et al. 2005a, b; Price et al. 2013). Fecundity decreased markedly with increasing numbers of oocysts in the inoculum. For example, a peak fecundity of $1.5 \times 10^5$ oocysts produced by a single oocyst at a dose of 100 oocysts was determined in the present study; fecundity decreased to only 163 oocysts from a single oocyst when birds were inoculated at $5 \times 10^4$ oocysts/bird. Oocyst output of *E. adenoeides* was lower than that of other species that infect the turkey such as *E. meleagrimitis* ($5.6 \times 10^3$ oocysts generated by a single oocyst at a dose of 100 oocysts) when compared under the same environmental conditions using the same strain and age of turkey pouls (El-Sherry et al. 2014). The simplest explanation of these observations is the crowding effect first described by Tyzzer et al. (1932) which is usually influenced by the tissue area available for the parasite to complete the endogenous development. Tyzzer demonstrated that the life cycle of cecal *Eimeria* spp., such as *E. tenella* in chickens, provides a smaller area of intestinal tissue to support endogenous development. Consequently, the crowding effect has a greater impact on cecal species, including *E. adenoeides*, compared with other species (e.g. *E. meleagrimitis*) that exploit almost the entire intestinal tract. Both direct damage to the cecal mucosa and sloughing of the cecal epithelium would decrease the cecal enterocytes available for further development and negatively affect total oocyst production. Crowding may also be responsible for the observation that younger (smaller) birds tend to shed fewer oocysts than older (larger) birds because of the formers’ comparatively smaller mucosa area.

**Acknowledgments** This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA), and from the Federal Economic Development Agency for Southern Ontario (FedDev) to J.R.B. and scholarship support from the Ministry of Higher Education and Research (MOHE), Egypt, to S.E.S. administered through the Bureau of Cultural and Educational Affairs (BCEA) of Egypt in Canada. Scholarship support by the Ontario Veterinary College (OVC PhD Scholarship) to M.E.O. is gratefully acknowledged. Julie Cobeam and Julia Whaley are thanked for their skilled technical assistance. The staff of the CAFA Animal Isolation Facility at the University of Guelph is thanked for their skilled care of experimental animals used in this study.

**References**


Chapman HD, Muthavarapu VK, Matsler PL, Chapman ME (2005a) Acquisition of immunity to *Eimeria maxima* in newly hatched chickens given 100 oocysts. Avian Dis 49:426–429


Hawkins PA (1952) Coccidiosis in turkey. Technical bulletin 226, East Lansing, MI: Michigan State College Agricultural Experimental Station


Long PL, Millard BJ, Shirley MW (1977) Strain variation within Eimeria meleagrimitis from the turkey. Parasitology 75:177–182


Williams RB (2001) Quantification of the crowding effect during infections with the seven Eimeria species of the domesticated fowl: its importance for experimental designs and the production of oocyst stocks. Int J Parasitol 31:1056–1069
APPENDIX 10: SEQUENCE-BASED GENOTYPING CLARIFIES CONFLICTING HISTORICAL MORPHOMETRIC AND BIOLOGICAL DATA FOR FIVE EIMERIA SPECIES INFECTING TURKEYS

(The contents of this chapter have been published as follows: El-Sherry, S, Ogedengbe, ME Hafeez, M Sayf-Al-Din, N Gad, JR Barta (2014). “Sequence-based genotyping clarifies conflicting historical morphometric and biological data for five Eimeria species infecting turkeys” Poultry Science 00:1–11 http://dx.doi.org/10.3382/ps/peu007
Sequence-based genotyping clarifies conflicting historical morphometric and biological data for 5 Eimeria species infecting turkeys

S. El-Sherry,1,2 M. E. Ogedengbe,1 M. A. Hafeez,1 M. Sayf-Al-Din,2 N. Gad,2 and J. R. Barta1∗

1Department of Pathobiology, Ontario Veterinary College, University of Guelph, Guelph, ON, N1G 2W1, Canada; and 2Department of Poultry Diseases, Faculty of Veterinary Medicine, Assiut University, Assiut, 71526, Egypt

ABSTRACT Unlike with Eimeria species infecting chickens, specific identification and nomenclature of Eimeria species infecting turkeys is complicated, and in the absence of molecular data, imprecise. In an attempt to reconcile contradictory data reported on oocyst morphometrics and biological descriptions of various Eimeria species infecting turkey, we established single oocyst derived lines of 5 important Eimeria species infecting turkeys, Eimeria meleagrimitis (USMN08–01 strain), Eimeria adenoeides (Guelph strain), Eimeria gallopavonis (Weybridge strain), Eimeria meleagridis (USAR97–01 strain), and Eimeria dispersa (Briston strain). Short portions (514 bp) of mitochondrial cytochrome c oxidase subunit I gene (mt COI) from each were amplified and sequenced. Comparison of these sequences showed sufficient species-specific sequence variation to recommend these short mt COI sequences as species-specific markers. Uniformity of oocyst features (dimensions and oocyst structure) of each pure line was observed. Additional morphological features of the oocysts of these species are described as useful for the microscopic differentiation of these Eimeria species. Combined molecular and morphometric data on these single species lines compared with the original species descriptions and more recent data have helped to clarify some confusing, and sometimes conflicting, features associated with these Eimeria spp. For example, these new data suggest that the KCH and KR strains of E. adenoeides reported previously represent 2 distinct species, E. adenoeides and E. meleagridis, respectively. Likewise, analysis of the Weybridge strain of E. adenoeides, which has long been used as a reference strain in various studies conducted on the pathogenicity of E. adenoeides, indicates that this coccidium is actually a strain of E. gallopavonis. We highly recommend mt COI sequence-based genotyping be incorporated into all studies using Eimeria spp. of turkeys to confirm species identifications and so that any resulting data can be associated correctly with a single named Eimeria species.

Key words: coccidiosis, oocyst, mitochondrial cytochrome c oxidase subunit I, turkey Eimeria, morphometrics

INTRODUCTION Coccidiosis is an important parasitic disease of turkey characterized by depression, anorexia, mucoid or bloody diarrhea, and growth retardation (Chapman, 2008). Seven species of protozoan parasites from the genus Eimeria are responsible for inducing coccidial infections in turkeys: Eimeria meleagrimitis, Eimeria dispersa, Eimeria innocua, Eimeria subrotunda, Eimeria meleagridis, Eimeria gallopavonis, and Eimeria adenoeides (Clarkson and Gentles, 1958). Although determining that turkeys are infected with Eimeria spp. is straightforward, identification of the specific parasite species involved is more complicated. This requires summation of information related to the prepatent period, morphology and morphometry of oocysts and other stages of the life cycle, site of development in the host, and macroscopic lesions induced by each species. These data are relatively nonspecific, and many features overlap among 2 or more Eimeria species. Finally, this analysis is too laborious to be used for routine diagnosis, and even then, it does not permit specific identification in many cases. Newer molecular techniques can allow more accurate diagnosis of Eimeria species, but suitable species-specific molecular markers must be identified (Schnitzler et al., 1998; Costa et al., 2001). Molecular characterization of Eimeria species using sequence-based genotyping (i.e., DNA barcoding) is an established molecular method for differentiating among Eimeria species. For coccidia, this process has been done using the nuclear (nu) 18S rDNA (e.g., Barta et al., 1997; Morrison, 2009; and many others), but there are issues with this particular genetic locus
because of the presence of paralogous rDNA loci within the nuclear genomes of at least some coccidia (e.g., Vrba et al., 2011; El-Sherry et al., 2013). In contrast, mitochondrial cytochrome c oxidase subunit I sequences (mt COI, 600 to 800 bp) have been shown to be highly effective in the differentiation of Eimeria species infecting a wide variety of vertebrates, including rabbits, rodents (Kvicerova and Hyspa, 2013), chickens (Ogedengbe et al., 2011; Vrba et al., 2011), and turkeys (El-Sherry et al., 2013, 2014a,b).

The growing turkey industry raises the need for accurate ways to diagnose and differentiate among Eimeria spp. infecting turkeys. Descriptions of Eimeria spp. infecting turkeys were done many years ago (1927 through 1958), and the quality and quantity of information included in each species description varied widely (Chapman, 2008). Creation of an effective and accurate molecular diagnostic procedure depends upon the availability of accurately identified DNA obtained from single oocyst line prepared for each Eimeria species to be included in the diagnostic method. The lack of consensus in the literature regarding the number and identity of Eimeria species infecting turkeys has limited the use of some PCR-based assays in routine diagnostic and epidemiological studies of coccidiosis in turkey.

In this study, we established and maintained 5 single oocyst derived lines of the most economically important species infecting turkeys, specifically E. meleagrimitis, E. adenoeides, E. gallopavonis, E. meleagridis, and E. dispersa. Each Eimeria species was characterized in detail to obtain information regarding the prepatent period, morphology and morphometry of oocysts and other stages of the life cycle, site of development in the host, and the macroscopic lesions they cause (e.g., El-Sherry et al., 2014a,b). Each was also characterized at the mt COI and nu 18S rDNA loci to provide reference sequence-based genotypes.

The objective of this study is to present linked morphometric and molecular data regarding 5 Eimeria spp. encountered typically in field samples obtained from turkeys. In the present work, uniformity of oocyst features (dimensions and oocyst structure) of each pure line was observed and suggested that revisions are required regarding the typical dimensions and morphological features for some of these parasites.

MATERIALS AND METHODS

Origins and Derivation of Single-Oocyst-Derived Lines of 5 Eimeria Species of Turkey

Eimeria meleagrimitis. A strain of E. meleagrimitis was obtained from a litter sample received by H. D. Chapman (Department of Poultry Science, University of Arkansas, Fayetteville) from a turkey farm near Willmar, Minnesota, on February 26, 2008. A pure line was derived by isolating a single oocyst using the method described by Shirley and Harvey (1996) and subsequently propagated in uninfected turkeys reared in an animal isolation unit at the University of Arkansas. This parent strain is referred to as E. meleagrimitis USMN08–01 and was kindly provided by H. D. Chapman on July 22, 2011. Four single-oocyst-derived lines of this parasite were generated by El-Sherry et al. (2013). One of the isolated lines (Line 4) was used for the present work. The biological characteristics of this strain and additional molecular markers were described by El-Sherry et al. (2014b) and Ogedengbe et al. (2014).

Eimeria adenoeides. The Guelph strain of E. adenoeides was obtained from a commercial turkey flock in Ontario, Canada, in approximately 1985. Four single-oocyst-derived lines of this parasite were generated by El-Sherry et al. (2013). One of the isolated lines (E. adenoeides Guelph Line 4) was used for the present work. The biological characteristics of this strain and additional molecular markers have been described by El-Sherry et al. (2014a) and Ogedengbe et al. (2014).

Eimeria gallopavonis. This strain was kindly provided by H. D. Chapman on July 22, 2011, under the name of “Weybridge strain of E. adenoeides” that was isolated in 1952 by S. F. M. Davies at the Central Veterinary Laboratory, Weybridge, United Kingdom. Oocysts of this parasite were obtained originally from pouls submitted for diagnosis following an outbreak of coccidiosis. This “Weybridge strain of E. adenoeides” was used subsequently in several studies (Hein, 1969; Joyner and Norton, 1972). The endogenous development and gross pathological lesions of this line of parasite (S. El-Sherry and J. R. Barta, unpublished observations) match the original species description of E. gallopavonis by Hawkins (1952) as does its large oocyst dimensions (see below). Consequently, this Weybridge isolate is named E. gallopavonis Weybridge strain in the present study. The complete mitochondrial genome sequence for this parasite has been published recently (Ogedengbe et al., 2014) and provides additional molecular markers for this parasite.

Eimeria meleagridis. Strain USAR97–01 of E. meleagridis was obtained in 1997 from the ceca of a turkey from a turkey farm in Northwest Arkansas and purified by H. D. Chapman (Department of Poultry Science, University of Arkansas, Fayetteville). This line was kindly provided by Chapman on July 22, 2011. Biological properties of this parent strain of E. meleagridis USAR97–01 have been reported recently (Matsler and Chapman, 2006), and additional molecular data on the single-oocyst-derived strain used in the present study have been recently reported by Ogedengbe et al. (2014).

Eimeria dispersa. The “Briston” strain of E. dispersa was isolated from turkeys from Briston, Norfolk, United Kingdom, by Long and Millard (1979). A sample of this strain was kindly provided by H. D. Chapman on July 22, 2011. Several single-oocyst lines were derived from each of these parent strains as described by Remmler and
McGregor (1964) with the modification that agar plugs carrying a single oocyst were given orally to the birds within gelatin capsules. For more details, see El-Sherry et al. (2013).

**Oocyst Measurements**

Measurements of 30 oocysts and 30 sporocysts of each species were made using a Provis AX70 photomicroscope (Olympus Canada, Richmond Hill, ON, Canada) fitted with a digital imaging device (Infinity3–1C, Lumenera Corporation Ottawa, ON, Canada) controlled using iSolution Lite image analysis software (Hoskin Scientific, Burlington, ON, Canada). All measurements were performed using a Provis AX70 photomicroscope (Olympus Canada, Richmond Hill, ON, Canada) fitted with a digital imaging device (Infinity3–1C, Lumenera Corporation Ottawa, ON, Canada) controlled using iSolution Lite image analysis software (Hoskin Scientific, Burlington, ON, Canada). All measurements were reported as means ± SD followed by range in parentheses.

**Molecular Characterization**

**DNA Extraction.** The DNA was isolated from oocysts as described by El-Sherry et al. (2013).

**PCR.** Short portions (514 bp) of mt COI for all of the 5 species were amplified using primers Eim‘COI366F-M13F (5'-TGTAAAACGACGGCCAG-TGGDTCWGGTRTWGGTTGGMC-3'; COI specific region underlined) and Eim‘COIF879R-M13R (5'-CAGGAAAACAGCTATGACCATATGRTGTGCCC-ADACT-3'; COI-specific region underlined) that were designed by examining an alignment of multiple sequences from various *Eimeria* species infecting poultry. The PCR reactions were carried out in an MJ Mini thermal cycler (Bio-Rad, Hercules, CA). The PCR reactions contained ~100 ng of genomic DNA from each *Eimeria* sp., 50 mM MgCl₂, 1 mM deoxynucleotide triphosphates (dNTP), 10 × PCR buffer, and 0.4 U Platinum Taq (Invitrogen, Burlington, ON, Canada). The PCR reaction thermal profile was as follows: initial heat activation of polymerase at 96°C for 10 min; 35 cycles of denaturation at 94°C for 30 s, annealing at 58°C for 30 s, extension at 72°C for 1 min; and a final extension at 72°C for 10 min. Both negative and positive template control reactions were included with each PCR run. The PCR products were electrophoresed on a 1.5% agarose submarine gel in 1 × Tris-acetate-EDTA buffer at 109 V for 45 min. The resulting gel was stained with ethidium bromide and the size of products estimated by comparison with a 100 bp to 10 kb DNA ladder (Bio Basic Inc., Mississauga, ON, Canada). A QIA quick gel extraction kit (Qiagen, Toronto ON, Canada) was used to purify excised bands. The resulting DNA was then sequenced in both directions with the forward and reverse amplification primers using an ABI Prism 7000 Sequence Detection System (Applied Biosystems Inc., Foster City, CA) by the Molecular Biology Unit of the Laboratory Services Division, University of Guelph (Guelph ON, Canada).

**Contig Generation, Sequence Alignment, and Phylogenetic Analyses**

Chromatograms from sequencing reactions were used to form high quality contigs with the de novo assembler in the Geneious bioinformatics software package (version 6.1 and later, available from http://geneious.com/); all PCR primer sites were identified in the resulting contigs and amino acid translations of the contig sequences using “Mold/Protozoan Mitochondrial” translation (translation table 4) were checked to confirm that the entire sequence from each parasite encoded the correct partial mt COI product. The resulting unambiguous sequences were aligned using Geneious and genetic distances calculated among all species.

Unrooted trees were generated from the aligned DNA sequences (fewer primers, 474 bp) using Bayesian inference, maximum parsimony (MP), and maximum likelihood (ML) tree-building methods [Mr. Bayes (Huelsenbeck and Ronquist, 2001), PAUP 4.0 (Swofford, 2003), and PhyML (Guindon and Gascuel, 2003), respectively], all implemented from within Geneious. In Bayesian inference and ML analyses, a GTR+1+gamma model of nucleotide substitution was used. The MP trees were generated using a branch and bound algorithm to find the optimal tree. Both ML and MP analyses were subjected to 100 bootstrap pseudoreplicates to generate a consensus tree for each of these tree-building algorithms to which bootstrap support values could be mapped.

**RESULTS**

**Oocyst Morphometrics**

A comparison of the morphometrics of sporulated oocysts and sporocysts (Figure 1) for the 5 *Eimeria* species is found in Table 1. Scatterplots illustrating the intraspecific variations of oocyst dimensions for individual species are found in Figure 2, and comparison of oocyst morphometrics among species is illustrated in Figure 3. Descriptions of the oocysts of the individual species follow.

**Eimeria meleagritis USMN08–01 Strain.** Sporulated oocysts (Figure 1a) were broadly ellipsoid in shape with a slight narrowing at one end. The oocyst wall was distinctly bi-layered and smooth. Micropyle and oocyst residuum were absent. A single refractile granule was typically detected at the one end of the oocyst that appeared slightly narrowed. Sporulated oocysts measured 18.9 ± 1.6 μm (17–22) by 15.7 ± 1.4 μm (13–19, n = 30) with a shape index (SI) of 1.21 ± 0.1. Sporocysts were ellipsoidal, 10.9 ± 0.4 μm (10–12) by 6.1 ± 0.7 μm (5–7, n = 30), and each possessed both a Stieda and sub-Stieda body (Figure 1b). Parastieda body was absent; sporocyst residuum was composed of granules of various sizes. Sporozoites possessed anterior and larger posterior refractile bodies.
Figure 1. Sporulated oocysts and sporocyst of 5 different Eimeria species infecting turkey: A, B) Sporulated oocyst and sporocyst of Eimeria melagrimitis (USMN08–01); C, D) Sporulated oocyst and sporocyst of Eimeria adenoeides (Guelph strain); E, F) Sporulated oocyst and sporocyst of Eimeria gallopavonis (Weybridge strain; note multiple refractile bodies); G, H) Sporulated oocyst and sporocyst of Eimeria meleagridis (USAR97–01); I, J) Sporulated oocyst and sporocyst of Eimeria dispersa (Briston strain; note the absence of refractile body). Scale bar = 10 μm.

Figure 2. Scatterplots of lengths (x-axis, μm) and widths (y-axis, μm; n = 30 per species) of 5 turkey Eimeria species: A) Eimeria melagrimitis (USMN08–01), B) Eimeria adenoeides (Guelph strain), C) Eimeria gallopavonis (Weybridge strain), D) Eimeria meleagridis (USAR97–01), and E) Eimeria dispersa (Briston strain). A morphometrically consistent population was observed for each species. Color version available online.
Table 1. Oocyst morphometrics of sporulated oocysts of Eimeria meleagrimitis, Eimeria adenoeides, Eimeria gallopavonis, Eimeria meleagridis, and Eimeria dispersa

<table>
<thead>
<tr>
<th>Species</th>
<th>Oocyst</th>
<th>Sporocyst</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Length (μm)</td>
<td>Width (μm)</td>
</tr>
<tr>
<td>Eimeria meleagrimitis (USMN08-01)</td>
<td>18.87 ± 1.56</td>
<td>15.66 ± 1.40</td>
</tr>
<tr>
<td></td>
<td>(15.6–22.1)</td>
<td>(13.1–18.8)</td>
</tr>
<tr>
<td>Eimeria adenoeides (Guelph strain)</td>
<td>18.69 ± 1.39</td>
<td>14.29 ± 1.01</td>
</tr>
<tr>
<td></td>
<td>(15.4–22.5)</td>
<td>(12.6–16.2)</td>
</tr>
<tr>
<td>Eimeria gallopavonis (Weybridge strain)</td>
<td>26.66 ± 1.93</td>
<td>18.62 ± 1.43</td>
</tr>
<tr>
<td></td>
<td>(23.6–31.1)</td>
<td>(14.6–20.6)</td>
</tr>
<tr>
<td>Eimeria meleagridis (USAR97-01)</td>
<td>26.28 ± 1.51</td>
<td>16.95 ± 1.21</td>
</tr>
<tr>
<td></td>
<td>(22.8–28.5)</td>
<td>(14.6–19.8)</td>
</tr>
<tr>
<td>Eimeria dispersa (Briston strain)</td>
<td>25.59 ± 1.12</td>
<td>20.95 ± 1.09</td>
</tr>
<tr>
<td></td>
<td>(23.8–27.9)</td>
<td>(19.4–23.2)</td>
</tr>
</tbody>
</table>

1Thirty oocysts and 30 sporocysts were measured from each species; all dimensions are expressed as a mean ± SD followed by the range in parentheses.

Figure 3. Diagrammatic summary of oocyst dimensions for 5 *Eimeria* species affecting turkeys. The mean lengths and widths (um) of 30 oocysts from each species are plotted, each surrounded by a dotted oval that represents 1 SD in the length and width means; the horizontal bars indicate the observed range of oocyst lengths and the vertical bars indicate the observed range of oocyst widths contributing to each mean. The dimensions of these oocysts are found in 2 groups: i) species possessing large oocysts (*E. meleagridis*, *E. gallopavonis*, and *E. dispersa*), and ii) species possessing small oocysts (*E. adenoeides and E. meleagrimitis*). Color version available online.
**Eimeria adenoecides Guelph Strain.** Sporulated oocysts (Figure 1c) were elliptoid in shape with a slight narrowing at one end. The oocyst wall was bi-layered and smooth. Micropyle and oocyst residuum were absent. A single refractile granule was typically detected at the narrow end. Sporulated oocysts measured 18.7 ± 1.4 μm (17–22) by 14.2 ± 0.9 μm (13–16, n = 30) with a SI of 1.31 ± 0.09. Sporocysts were ellipsoidal, measured 9.7 ± 0.7 μm (9–11) by 5.9 ± 0.5 μm (4–7, n = 30), and each possessed a Stieda and sub-Stieda body (Figure 1d). Parastieda body was absent; sporocyst residuum was composed of granules of various sizes. Sporozoites possessed anterior and larger posterior refractile bodies.

**Eimeria gallopavonis Weybridge Strain.** Sporulated oocysts (Figure 1e) were broadly ovoid with a consistent narrowing at one end. The double-layered oocyst wall was smooth. Micropyle and oocyst residuum were absent. Multiple refractile granules (up to 4 granules) were detected attached to the inner layer of the oocyst wall at the narrow end. Sporulated oocysts measured 26.7 ± 1.9 μm (24–31) by 18.6 ± 1.4 μm (16–21, n = 30) with a SI of 1.44 ± 0.10. Sporocysts were ellipsoidal, 12.3 ± 0.6 μm (11–13) by 7.7 ± 0.5 μm (7–9, n = 30), and each possessed both a Stieda and sub-Stieda body (Figure 1f). Parastieda body was absent; sporocyst residuum was composed of granules of various sizes. Sporozoites possessed anterior and larger posterior refractile bodies.

**Eimeria meleagridis USAR97–01 Strain.** Sporulated oocysts (Figure 1g) were ovoid with a consistent narrowing at one end. The double-layered oocyst wall was smooth. Micropyle and oocyst residuum were absent. A single refractile polar granule could be detected attached to the inner layer at the narrow end. Sporulated oocysts measured 26.3 ± 1.5 μm (23–28) by 16.9 ± 1.2 μm (15–20, n = 30) with SI of 1.56 ± 0.11. Sporocysts were ellipsoidal, 11.2 ± 0.5 μm (10–12) by 6.2 ± 0.3 μm (6–7, n = 30), and each possessed both a Stieda and sub-Stieda body (Figure 1h). Parastieda body was absent; sporocyst residuum was composed of granules of various sizes. Sporozoites possessed anterior and larger posterior refractile bodies.

**Eimeria dispersa Briston Strain.** Sporulated oocysts (Figure 1i) were subspherical in shape. The double-walled oocyst was smooth. Micropyle and oocyst residuum are absent. Refractile polar granule was absent. Sporulated oocysts measured 25.6 ± 1.1 μm (24–28) by 20.9 ± 1.1 μm (19–23, n = 30) with SI of 1.22 ± 0.06. Sporocysts were ellipsoidal, 14.4 ± 1.0 μm (12–16) by 8.2 ± 0.6 μm (7–9, n = 30), and each possessed both a Stieda and sub-Stieda body (Figure 1j). Parastieda body was absent; sporocyst residuum was composed of granules of various sizes. Sporozoites possessed anterior and larger posterior refractile bodies.

### Mitochondrial Cytochrome C Oxidase Subunit I Partial Gene Sequences

The 366F/679R primer set successfully amplified and primed the sequencing of partial mt COI sequences from *E. meleagrimitis* USMN08–01 (KJ526131), *E. adenoeides* Guelph strain (KJ526134), *E. gallopavonis* Weybridge strain (KJ526132), *E. meleagridis* USAR97–01 strain (KJ526133), and *E. dispersa* Briston strain (KJ526130). Examination of the sequence assemblies demonstrated no intraspecific variation in any of the PCR fragments that were sequenced.

Comparison of mt COI sequences from these 5 *Eimeria* species of turkey showed sufficient species-specific sequence variation to support short mt COI sequences as species-specific markers (Table 2). Pairwise single nucleotide differences (SND) between *Eimeria* spp. ranged from 15 to 79 SND; the largest variation between 2 *Eimeria* spp. infecting turkeys was between *E. meleagrimitis* and *E. dispersa* (79 SND, 84.6% sequence identity), and the interspecific genetic distance was the lowest between *E. meleagridis* and *E. adenoeides* (15 SND, 97.1% sequence identity).

### Phylogenetic Analyses Using mt COI Sequences

The unrooted tree generated using MrBayes for mt COI sequences from the 5 different species of *Eimeria* infecting turkeys is illustrated in Figure 4. Maximum parsimony and maximum likelihood trees (trees not shown) had identical branching orders. For mt COI sequences analysis, sequences form 2 main clades, one of them composed of the 3 cecal species (*E. adenoeides, E. gallopavonis*, and *E. meleagridis*) and the other containing the 2 species (*E. dispersa* and *E. meleagrimitis*) that do not infect the ceca of turkeys.
DISCUSSION

The first *Eimeria* species formally described from turkeys were named by Tyzzer (1929) when he published brief descriptions of *E. meleagridis* (ellipsoidal, 23.7 by 17.38 μm with 1 or 2 polar granules), *E. meleagrimitis* (18 by 15.25 μm, with a single globular polar granule), and *E. dispersa* (22.75 by 18.84 μm, ovoidal without any polar granules). Tyzzer (1929) was able to transmit *E. dispersa* to turkeys from the eastern bobwhite quail. Hawkins (1952) described *E. dispersa* oocysts from turkey strain to be larger (26.07 by 21.04 μm) but with the same morphological features reported by Tyzzer (1929) for the type material isolated from bobwhite quail. Hawkins (1952) described a new species, *E. gallopavonis*, with the largest oocyst of any coccidium described previously from turkeys. Concurrently, Moore and Brown (1951, 1952) and Moore et al. (1954) described 3 additional species, *E. adenoeides*, *E. subrotunda*, and *E. innocua*.

Careful examination of the original species descriptions makes it clear that accuracy of the reported oocyst dimensions (Table 3) may have been affected by the purity of the isolate upon which the description of each species was based. For example, in the species description of *Eimeria adenoeides* in North American turkeys by Moore and Brown (1951), the authors recorded a tremendously wide range of dimensions for their new species (18.9 to 31.3 μm in length; 12.6 to 20.9 μm in width). By comparison, none of our purified lines of turkey coccidia had a range of oocyst lengths exceeding 7 μm within a single line. Although it is possible that Moore and Brown’s (1951) new species possessed highly pleomorphic oocysts, the more likely explanation is the presence of more than one *Eimeria* species in the sample. The micrograph in the original species description (Figure 1 of Moore and Brown, 1951) for oocysts in their work revealed that there are at least 3 apparent oocyst morphotypes present in the sample that was used to describe *E. adenoeides*. Later, in Clarkson’s (1958) characterization of a strain of an *Eimeria* sp. that was isolated in England that he identified as *E. adenoeides*, multiple oocyst morphotypes were apparent in the illustration of the sporulated oocysts (Figure 4 of Clarkson, 1958) and the histogram of oocyst lengths and widths (Figure 5 of Clarkson, 1958). Clarkson (1958) even noted 2 apparently different prepatent periods resulting from infections with this isolate. In addition, several authors have explained variation of oocyst morphometrics in isolates as different strains of the same species; this
Table 3. Comparison of oocyst morphometrics reported for 5 Eimeria species infecting turkeys

<table>
<thead>
<tr>
<th>Item</th>
<th>Length (μm)</th>
<th>Width (μm)</th>
<th>Average (μm)</th>
<th>Shape index</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eimeria meleagrimitis</td>
<td>18</td>
<td>15.25</td>
<td>18 × 15.25</td>
<td>1.18, subspherical</td>
<td>Tyzzer, 1929</td>
</tr>
<tr>
<td></td>
<td>15.8–20.9</td>
<td>13.1–21.9</td>
<td>19.1 × 16.2</td>
<td>1.17, subspherical</td>
<td>Hawkins, 1952</td>
</tr>
<tr>
<td>KR strain</td>
<td>20.3</td>
<td>16.4</td>
<td>20.3 × 16.4</td>
<td>1.24</td>
<td>Vrba and Pakandl, 2014</td>
</tr>
<tr>
<td>USMN08-01 strain</td>
<td>17–22</td>
<td>13–19</td>
<td>18.9 × 15.7</td>
<td>1.21 ± 0.10, broadly elliptoid</td>
<td>Present study</td>
</tr>
<tr>
<td>Eimeria adenoeides</td>
<td>18.96–31.26</td>
<td>12.6–20.94</td>
<td>25.6 × 16.6</td>
<td>1.54, ellipsoidal</td>
<td>Moore and Brown, 1951</td>
</tr>
<tr>
<td>“E. meleagris” KCH strain</td>
<td>18.1</td>
<td>14.8</td>
<td>18.1 × 14.8</td>
<td>1.22, ovoid</td>
<td>Vrba and Pakandl, 2014</td>
</tr>
<tr>
<td>“E. meleagris” NR strain</td>
<td>21.8</td>
<td>15.3</td>
<td>21.8 × 15.3</td>
<td>1.43, ellipsoidal</td>
<td>Vrba and Pakandl, 2014</td>
</tr>
<tr>
<td>Guelph strain</td>
<td>17–22</td>
<td>13–16</td>
<td>18.7 × 14.2</td>
<td>1.31 ± 0.09, elliptoid</td>
<td>Present study</td>
</tr>
<tr>
<td>Eimeria gallopavonis</td>
<td>22.2–32.7</td>
<td>15.2–19.4</td>
<td>27.1 × 17.2</td>
<td>1.52, long ellipsoidal</td>
<td>Hawkins, 1952</td>
</tr>
<tr>
<td>KR strain</td>
<td>26.6</td>
<td>16.4</td>
<td>26.6 × 16.4</td>
<td>1.62</td>
<td>Vrba and Pakandl, 2014</td>
</tr>
<tr>
<td>Weybridge strain</td>
<td>24–31</td>
<td>16–21</td>
<td>26.7 × 18.6</td>
<td>1.44 ± 0.10, broadly ovoid</td>
<td>Present study</td>
</tr>
<tr>
<td>Eimeria meleagrimitis</td>
<td>23.79</td>
<td>17.38</td>
<td>23.9 × 17.38</td>
<td>1.38, ellipsoidal</td>
<td>Tyzzer, 1929</td>
</tr>
<tr>
<td></td>
<td>20.3–30.8</td>
<td>15.4–20.6</td>
<td>24.4 × 18.1</td>
<td>1.35, ellipsoidal</td>
<td>Hawkins, 1952</td>
</tr>
<tr>
<td></td>
<td>“E. adenoeides” KR strain</td>
<td>18.8–34</td>
<td>14.1–23.5</td>
<td>27.8 × 19.4</td>
<td>1.43, ellipsoidal</td>
</tr>
<tr>
<td></td>
<td>22.8</td>
<td>14.9</td>
<td>22.8 × 14.9</td>
<td>1.53, ellipsoidal</td>
<td>Vrba and Pakandl, 2014</td>
</tr>
<tr>
<td></td>
<td>USAR97–01 strain</td>
<td>23–28</td>
<td>15–20</td>
<td>26.3 × 16.9</td>
<td>1.56 ± 0.11, ovoid</td>
</tr>
<tr>
<td>Eimeria dispersa</td>
<td>22.75</td>
<td>18.84</td>
<td>22.79 × 18.84</td>
<td>1.21, ovoidal</td>
<td>Tyzzer, 1929</td>
</tr>
<tr>
<td>Briston strain</td>
<td>24–28</td>
<td>19–23</td>
<td>25.6 × 20.9</td>
<td>1.22 ± 0.06, subspherical</td>
<td>Present study</td>
</tr>
</tbody>
</table>

practice has complicated the use of oocyst morphometrics for species identifications. Early work of Long et al. (1977) described an E. meleagris strain that was smaller (16.2 μm by 14.8 μm versus 18 μm by 15.25 μm) than what was originally reported by Tyzzer (1929) and Hawkins (1952). More recently, 2 strains of E. adenoeides were reported with marked variation (up to 43% variability in the length) by Poplstein and Vrba (2011). The question remains as to whether there is such dramatic morphological variation among strains of individual species or whether these “strains” are in reality distinct species that have been mistakenly assigned the same species name. The present study compared only 5 single-oocyst-derived isolates representing 5 Eimeria species that infect turkeys. It is highly likely that isolates from other geographic regions of one or more of these parasites may demonstrate morphological variation from the lines that we examined herein. Indeed, a recent report on 3 isolates of Eimeria adenoeides (strains KCH, RM, and NR identified as E. meleagris according to Vrba and Pakandl, 2014) that share COI sequence identity with E. adenoeides Guelph strain reported in the present paper have somewhat variable oocyst shape indices of 1.22, 1.40, 1.43, and 1.31, respectively.

Sequence-based genotyping can help to resolve the conflicting morphometric data. DNA barcoding using partial mt COI genes sequences was effective for differentiation of Eimeria species infecting chickens (Ogedengbe et al., 2011; Vrba et al., 2011) and was a stable genetic marker for 2 Eimeria species infecting turkeys even when the more commonly used genetic locus, nu 18S rDNA, was comparatively unreliable (El-Sherry et al., 2013). The primers that were designed and used in the present study successfully produced short sequences (~500 bp) from each Eimeria species that can be used not only in identification and differentiation of Eimeria spp. in turkey, but also can be used as a powerful tool to check the purity of each line during subsequent propagation. For example, despite somewhat variable oocyst shape indices, the 4 lines of E. adenoeides discussed previously (strains KCH, Guelph, NR, RM; Vrba and Pakandl, 2014 and present study) have identical COI sequences. Eimeria meleagrimitis KR (Vrba and Pakandl, 2014) and E. meleagris USAR97–01 (present study) have identical COI sequences and near-identical oocyst shape indices of 1.53 and 1.56, respectively.

A diagrammatic representation of the mean lengths and widths of oocysts, as well as the SD and ranges of dimensions observed, from each of the 5 species examined in the present study demonstrates the largely overlapping dimensions of these parasites (Figure 4). The oocysts fall into 2 groups based solely on dimensions: 1) species possessing large oocysts (E. meleagris, E. gallopavonis, and E. dispersa), and 2) species possessing small oocysts (E. adenoeides and E. meleagris). Of the 3 characterized parasites with large oocysts, E. gallopavonis is reported to be the largest (27.1 × 17.23 μm, Hawkins, 1952). The Weybridge strain of E. gallopavonis used in the present study was isolated in 1952 and subsequently used in several research papers (Hein, 1969 on pathogenicity; Joyner and Norton, 1972 on drug sensitivity); however, in both papers the strain was named E. adenoeides (Weybridge strain). The oocysts of this strain are larger than dimensions reported originally for E. adenoeides by Moore and Brown (1951). The endogenous development of E. gallopavonis Weybridge strain used in the present study (S. El-Sherry and J. R. Barta, unpublished observations) matches the original species description by Hawkins (1952) and its large oocyst dimensions match those of original species description as well (26.66 ± 1.93 μm by 18.62 ± 1.43 μm). These findings indicate that Hein (1969) and Joyner and Norton (1972) were
likely working with *E. gallopavonis* and that their conclusions should be interpreted in that light. Like *E. gallopavonis*, both *E. meleagridis* and *E. dispersa* have large oocysts (26.28 ± 1.51 μm by 16.95 ± 1.21 μm for *E. meleagridis* and 25.59 ± 1.12 μm by 20.95 ± 1.09 μm for *E. dispersa*). However, these 3 species can be differentiated microscopically by the presence or absence of polar granules. *Eimeria gallopavonis* (Weybridge) oocysts contain up to 4 refractile globules that appear attached to the inner oocyst wall at its narrow end. The oocysts of *E. meleagridis* (USAR97–01 strain) contain only a single refractile body at the narrow end and the oocysts of *E. dispersa* do not contain any refractile body whatsoever. Additionally, the oocysts of *E. dispersa* are more spherical in shape (SI = 1.22 ± 0.06) than either *E. meleagridis* (SI = 1.56 ± 0.11) or *E. gallopavonis* (SI = 1.44 ± 0.10).

Deciding on what are typical dimensions and morphological features for oocysts of *E. adenoeides* was especially challenging. It is highly likely that Moore and Brown (1951), and likely the later researchers as well, were working with mixed samples containing multiple species. Detailed examination of the endogenous development of *E. adenoeides* Guelph strain (El-Sherry et al., 2014a) has shown that this parasite develops similarly to the original species description of *E. adenoeides* by Moore and Brown (1951). Recently, Popelstov and Vrba (2011) isolated 2 strains of coccidia from turkeys, *Eimeria adenoeides* (KR strain) had large ellipsoidial oocysts measuring ~27.8 × 19.4 μm, whereas *E. adenoeides* (KCH strain) had smaller oval oocysts measuring ~16.1 × 19.4 μm. Comparing the mt COI sequences and nu 18S sequences of the KCH and KR strains of *E. adenoeides* with our single-oocyst derived strains showed that the Guelph strain of *Eimeria adenoeides* matches the *E. adenoeides* KCH strain genetically. In contrast, the *E. adenoeides* KR strain COI and 18S sequences match the sequences we obtained from a single-oocyst line of *E. meleagridis* (derived from the USAR97–01 strain) and the oocyst dimensions of these latter strains likewise agree. To add further confusion, Vrba and Pakandl (2014) recently reassigned their *E. adenoeides* KCH and KR strains, as well as strains RM and NR, to the single species *E. meleagridis*. The *E. meleagridis* KR strain and USAR97–01 strains are morphologically (Vrba and Pakandl, 2014; present study) and genetically (Vrba and Pakandl, 2014; Ogedengbe et al., 2014) linked and should be considered a single species but distinct from the other 4 strains (KCH, RM, NR, and Guelph). We continue to refer to these latter strains as *E. adenoeides*.

The small oocysts of *E. adenoeides* Guelph strain cannot be differentiated morphometrically from the oocysts of *E. meleagrimitis* (USMN08–01 strain) because they overlap substantially in lengths, widths and morphology (Figure 4). In contrast, the sites of endogenous development and pathological lesions induced by *E. adenoeides* and *E. meleagrimitis* are completely different. Lesions of *E. meleagrimitis* are located mainly in the upper and middle part of the intestinal tract and never in the cecum (El-Sherry et al., 2014b). In the original description of *E. adenoeides* by Moore and Brown (1951), lesions were described in the lower intestinal tract and, notably, within the cecum with the formation of a characteristic caseous cecal core. Formation of a characteristic corrugated caseous plug within the ceca of infected turkeys by *E. adenoeides* Guelph strain was observed consistently in experimental infections (El-Sherry et al., 2014a). Contamination of *E. meleagrimitis* cultures with small numbers of morphologically indistinguishable *E. adenoeides* oocysts may explain why the *E. meleagrimitis* lifecycle was frequently reported to include some endogenous development within the cecal pouche (i.e., Hawkins, 1952 and Clarkson, 1959).

The severity and the location of the pathological lesions of *E. adenoeides* can be confused with *E. meleagridis* and *E. gallopavonis*. Lesions associated with *E. gallopavonis* are usually confined to the lower intestinal tract distal to Meckel’s diverticulum within the lower ileum, cecal necks, and rectum, but sparing the cecal pouches (Farr et al., 1961; Wehr et al., 1962; Lund and Farr, 1965). The strain used in the present study induces the same characteristic pathological lesions described by Hawkins (1952) for *E. gallopavonis* (S. El-Sherry and J. R. Barta, unpublished observations). Hein (1969) described identical lesion formation for this same *E. gallopavonis* Weybridge strain but under the name of *E. adenoeides* (Weybridge strain).

The pathological lesions associated with *E. meleagridis* are reported to be confined to the cecum by Hawkins (1952) and Clarkson (1959) who reported the formation of caseated plug in the cecum. However, work describing a recently isolated strain of *E. meleagridis* USAR97–01 strain by Matsler and Chapman (2006) did not demonstrate typical caseous cores at any location in the digestive tract; in the present study, lines derived from this same strain were capable of producing marked lesions in the cecal pouch with a characteristic cecal core located mainly in the distal part of the cecal pouch (S. El-Sherry and J. R. Barta, unpublished observations). The observed variations in macroscopic lesion formation and pathogenicity of this line likely reflect differences in the experimental methods used (e.g., type [breed] and age of birds, type of feed, freshness of the oocyst culture, inoculating dose, etc.).

Unlike with *Eimeria* species infecting chickens (e.g., Reid and Long, 1979), specific identification and nomenclature of *Eimeria* species infecting turkeys is complicated and, in the absence of molecular data, imprecise. The inability to specifically identify the coccidia that were used to generate data on the life cycles, endogenous development, oocyst dimensions, and pathogenicity of various *Eimeria* “species” over the last 8 decades has resulted in a confused, and often contradictory data set (summarized by Chapman, 2008).

Use of molecular markers (specifically sequence-based genotyping using the mt COI locus) has permitted the unequivocal linkage of oocyst morphometrics and
endogenous development with a highly discriminating genetic marker. The stability of this marker is evident from identical, or almost identical, partial mt COI sequences possessed by various lines of coccidia of within each of 5 Eimeria spp. infecting turkeys that were derived independently in Europe (see Vrba and Pakandl, 2014) and in North America (Ogedengbe et al., 2014; present study). Intraspecific variation at the COI locus ranged from 0 to 0.2% (3 SND over 1,257 bp), whereas interspecific variation ranged from 2.6% (33 SND) to 15.2% (191 SND; data not shown); these data strongly support use of the COI locus for genotyping of Eimeria species.

Much of the early literature describing Eimeria species infecting turkeys and the effects these parasites have on their hosts is largely uninformative because the identity(ies) of the organism(s) under study cannot be determined with any confidence and this reflects limitations of the differentiating criteria used at the time. Genotyping (e.g., mt COI sequence-based typing) should be used to identify unambiguously any Eimeria species before its use in studies on oocyst morphology, endogenous development, or pathogenicity of these parasites. Such genotyping of these parasites will help to ensure the validity of comparisons among studies originating from various research groups and, perhaps, may permit researchers to reliably document the range of phenotypic variation that can be expressed by a single Eimeria species.

ACKNOWLEDGMENTS

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada, the Ontario Ministry of Agriculture, Food and Rural Affairs, and from the Federal Economic Development Agency for Southern Ontario to J.R.B. and scholarship support from the Ministry of Higher Education and Research, Egypt, administered through the Bureau of Cultural and Educational Affairs of Egypt in Canada to S.E.S. H. David Chapman, University of Arkansas, Fayetteville, is thanked for providing Eimeria meleagrimitis, Eimeria meleagridis, Eimeria gallopavonis, and Eimeria dispersa isolates from which we derived the single-oocyst lines used in this study. Careful care of our research animals by the staff of the Research Isolation Unit of the Central Animal Facility, University of Guelph, is greatly appreciated. Technical assistance of Julie Cobean and Julia Whale (Pathobiology, University of Guelph) is greatly appreciated.

REFERENCES


APPENDIX 11: SIMULTANEOUS IDENTIFICATION AND DNA BARCODING OF 6
*Eimeria* SPECIES INFECTING TURKEYS USING PCR PRIMERS TARGETING THE
MITOCHONDRIAL CYTOCHROME C OXIDASE SUBUNIT-1 (mtCOI) LOCUS

(The contents of this chapter have been published as follows: Mian A. Hafeez, Srichaitanya Shivaramaiah, Kristi M. Dorsey, Mosun E. Ogedengbe, Shiem El-Sherry, Julia Whale, Julie Cobean and John R. Barta (2014).
“Simultaneous identification and DNA barcoding of *Eimeria* species infecting turkeys using PCR primers targeting the mitochondrial cytochrome c oxidase subunit-1 (mtCOI) locus”. Parasitology Research doi 10.1007/s00436-015-4361-y)
Simultaneous identification and DNA barcoding of six \textit{Eimeria} species infecting turkeys using PCR primers targeting the mitochondrial cytochrome \textit{c} oxidase subunit I (mtCOI) locus

Mian A. Hafeez \cdot Srichaitanya Shivaramaiah \cdot Kristi Moore Dorsey \cdot Mosun E. Ogedengbe \cdot Shiem El-Sherry \cdot Julia Whale \cdot Julie Cobean \cdot John R. Barta

Received: 12 September 2014 / Accepted: 30 January 2015 © Springer-Verlag Berlin Heidelberg 2015

Abstract Species-specific PCR primers targeting the mitochondrial cytochrome \textit{c} oxidase subunit I (mtCOI) locus were generated that allow for the specific identification of the most common \textit{Eimeria} species infecting turkeys (i.e., \textit{Eimeria adenoeides}, \textit{Eimeria meleagrimitis}, \textit{Eimeria gallopavonis}, \textit{Eimeria meleagris}, \textit{Eimeria dispersa}, and \textit{Eimeria innocua}). PCR reaction chemistries were optimized with respect to divalent cation (MgCl\textsubscript{2}) and dNTP concentrations, as well as PCR cycling conditions (particularly anneal temperature for primers). Genomic DNA samples from single oocyst-derived lines of six \textit{Eimeria} species were tested to establish specificity and sensitivity of these newly designed primer pairs. A mixed 60-ng total DNA sample containing 10 ng of each of the six \textit{Eimeria} species was used as DNA template to demonstrate specific amplification of the correct product using each of the species-specific primer pairs. Ten nanograms of each of the five non-target \textit{Eimeria} species was pooled to provide a non-target, control DNA sample suitable to test the specificity of each primer pair. The amplifications of the COI region with species-specific primer pairs from pooled samples yielded products of expected sizes (209 to 1,012 bp) and no amplification of non-target \textit{Eimeria} sp. DNA was detected using the non-target, control DNA samples. These primer pairs specific for \textit{Eimeria} spp. of turkeys did not amplify any of the seven \textit{Eimeria} species infecting chickens. The newly developed PCR primers can be used as a diagnostic tool capable of specifically identifying six turkey \textit{Eimeria} species; additionally, sequencing of the PCR amplification products yields sequence-based genotyping data suitable for identification and molecular phylogenetics.

Keywords Molecular characterization \cdot Turkey coccidia \cdot Diagnostics \cdot Phylogenetic analysis \cdot Species identification

Introduction

Coccidiosis is a ubiquitous and economically important threat to the poultry industry (Long et al. 1977). Coccidiosis may cost the broiler chicken industry alone over USD$ 3 billion annually with almost 70% of this estimated cost is due to subclinical coccidiosis (Allen and Fetterer 2002; Long 1973; Dezfulian et al. 2010). The disease is caused by protozoan parasites belonging to the genus \textit{Eimeria}. In the domestic turkey (\textit{Meleagris gallopavo}), seven species have been described: \textit{Eimeria adenoeides}, \textit{Eimeria meleagrimitis}, \textit{Eimeria gallopavonis}, \textit{Eimeria meleagris}, \textit{Eimeria dispersa}, \textit{Eimeria innocua}, and \textit{Eimeria subrotunda} (McDougall 2003; Chapman 2008). Four species (\textit{E. adenoeides}, \textit{E. meleagrimitis}, \textit{E. gallopavonis}, and \textit{E. dispersa}) are of primary economic concern due to morbidity and mortality experienced in commercial turkeys associated with these species (Cook et al. 2010). Coccidiosis is a disease complex with
simultaneous infections common and with each *Eimeria* species making separate, distinct contributions to the pathogenicity in the host. Specific diagnosis of the agent(s) responsible for coccidiosis in turkeys can be challenging due to similar oocyst dimensions, nonspecific lesion development, and overlapping biological features among the potential causative *Eimeria* species (Long et al. 1977; Chapman 2008; El-Sherry et al. 2014a, 2015).

Species identification has been based on morphological features of oocysts (shape, size, and refractive granules), biological features (e.g., pre-patent period or sporulation time), or pathological changes induced by the parasites (e.g., intestinal site and shape of lesion) similar to those criteria used to differentiate *Eimeria* species infecting chickens (Joyner and Long 1974). Mixed infections can make precise identification based on lesions in an infected intestinal tract nearly impossible (Jeffers 1975; McDougald and Jeffers 1976). Control of coccidiosis is based on prophylactic use of anticoccidial drugs in the feed but live vaccines are being used increasingly in the field (Chapman et al. 2002; Allen and Fetterer 2002).

Reliable species discrimination is imperative for vaccine development and production as well as epizootiological and population biology studies. In the past, different tests have been used to identify the pathogenic *Eimeria* species infecting turkeys. The earliest methods of species differentiation were based on species-specific immune responses.

Molecular differentiation of species was first attempted using isoenzyme analysis for species identification through molecular polymorphism (Shirley 1975; Johnston and Fernando 1997). Thereafter, PCR assays were described for identification of *Eimeria* species using different regions of the ribosomal cistrons including 5S rDNA, small subunit (18S) rDNA, and the ribosomal internal transcribed spacer regions 1 and 2 (ITS-1 and ITS-2, respectively) (Stucki et al. 1993; Tsuji et al. 1997; Schnitzler et al. 1998; Woods et al. 2000; Gasser et al. 2001; Su et al. 2003; Morrison et al. 2004; Rampin et al. 2006; Cook et al. 2010; Poplstein and Vrba 2011; Ogedengbe et al. 2011; El-Sherry et al. 2013). SCAR-based, species-specific PCR assays were developed for *Eimeria* spp. infecting the domestic fowl (Fernandez et al. 2003a, b). However, only a single PCR-based method for the specific identification of various *Eimeria* species in turkeys has been published (Cook et al. 2010). This assay targeted the ITS-1 region of the ribosomal gene array and could identify only four species specifically. The nuclear rDNA is subject to recombination events and, in at least some coccidia, 18S rDNA (and presumably the adjoining ITS-1 region) has been duplicated within the genome and the paralogs are highly divergent (e.g., Vrba et al. 2011; El-Sherry et al. 2013).

The mitochondrial cytochrome *c* oxidase subunit I (mtCOI) locus is present in many apicomplexan parasites and has been used extensively for sequence-based genotyping and species identification (so-called DNA barcoding) of numerous organisms that exploit oxidative phosphorylation as an energy source (Hebert et al. 2003). The mitochondrial genome has the advantage of replicating mitotically with strict maternal inheritance making recombination events less likely. For these reasons, genes encoded in the mitochondrial genome have also been used extensively in molecular phylogenetics of many organisms, including haemosporidian parasites (Perkins and Schall 2002), but these loci have not been exploited widely for inferring the evolutionary history of coccidian parasites (Traversa et al. 2007). Sequences from mtCOI of a number of *Eimeria* species have been demonstrated to be much more robust than nu 18S rDNA for identification and differentiation of closely related parasites (e.g., Ogedengbe et al. 2011). The mtCOI locus appears to lack the paralog issues demonstrated in the rDNA of these parasites (El-Sherry et al. 2013).

The nature of the mtCOI locus of coccidia has been exploited in this paper to generate a set of species-specific PCR primers based on this locus that permit the specific identification of six common *Eimeria* species infecting the domestic turkey and, by optional direct sequencing of the resulting diagnostic PCR fragments, DNA barcode data that can be used for molecular phylogenetics.

### Materials and methods

#### Parasites

Numerous, single oocyst-derived lines of five *Eimeria* species infecting domestic turkeys were available for this study. Each line with the origin of the original isolates from which each was derived is summarized by El-Sherry et al. (2015). The parasites available for study are as follows: *E. adenoeides* Guelpf strain (El-Sherry et al. 2014b, 2015), *E. meleagrimitis* USMN08-01 strain (see Long et al. 1977; El-Sherry et al. 2014a, 2015 for biological features), *E. gallopavonis* Weybridge strain (see Hein, 1969 for biological features), *E. gallopavonis* USKS06-01, *E. meleagrisidis* USAR97-01 strain (see Matsler and Chapman 2006 for biological features), and *E. dispersa* Briston strain.

All animal experimentation was conducted in coccidia-free birds in the Animal Isolation Unit of the Campus Animal Facility (University of Guelph, Guelph ON, Canada); all experimental procedures were reviewed and approved by the University Animal care Committee and complied with the Canadian Council on Use of Experimental Animals (2nd edition).

#### DNA extraction from oocysts

Isolation of DNA was performed after purification and surface treatment to reduce fecal contamination. Partially purified oocysts were suspended in saturated salt solution (sat. NaCl,
PCR amplification was accomplished using six newly designed species-specific primer pairs for *Eimeria* spp. infecting turkeys. These primer pairs, recommended PCR annealing temperatures, and expected PCR product sizes are found in Table 1.

PCR reactions were carried out in an MJ Mini® thermal cycler (Bio-Rad Laboratories (Canada) Ltd., Mississauga, ON, Canada) using 50 ng of template DNA, 50 mM MgCl₂, 1 mM dNTPs, 1× PCR buffer, and 0.4 U Platinum® Taq polymerase (Life Technologies Inc., Burlington ON, Canada). The PCR thermal profile was as follows: initial heat activation of polymerase at 95 °C for 10 min; 35 cycles of denaturation at 94 °C for 30 s, annealing at 50–62 °C for 45 s (anneal temperatures vary depending on the species-specific primer pair) and extension at 72 °C for 70 s; and a final extension at 72 °C for 10 min. Both negative and positive template control reactions were included with each PCR run. PCR products were electrophoresed on a 1.5 % agarose submarine gel in Tris-acetate-EDTA (TAE) buffer (40 mM Tris, 20 mM acetic acid, and 1 mM EDTA) at ~100 V for ~45 min. The resulting gel was stained with ethidium bromide and the product size estimated by comparison with a 100-bp to 10-kb DNA ladder (Bio Basic Inc., Mississauga ON, Canada) visualized using UV transillumination.

Optimization of PCR for specificity and sensitivity

Various equimolar mixtures of genomic DNA samples obtained from single oocyst-derived lines of six *Eimeria* species infecting domestic turkeys were prepared to permit testing of the specificity and sensitivity of these newly designed primer pairs. A mixed DNA sample containing equivalent amounts of each of the six *Eimeria* species was used as the DNA template to demonstrate that PCR could amplify specifically the correct product using each of the species-specific primer pairs from a mixed *Eimeria* spp. template. In such reactions, 60 ng total DNA (10 ng per *Eimeria* sp.) was used as template for each of the PCR primer pairs individually. To test for the specificity of each of the primer pairs, equivalent amounts of each of the five non-target *Eimeria* spp. were mixed to provide a negative control DNA sample suitable for each primer pair. For example, the mixed species negative template DNA to test the specificity of the *E. adenoideides*-specific primers would contain 10 ng of genomic DNA from each of *E. meleagrimitis, E. dispersa, E. meleagris*, *E. gallopavonis*, and *E. innocua*. A total of six “non-target” negative control DNA samples were made. Similarly, a mixture containing equal amounts of genomic DNA from the seven *Eimeria* species infecting the domestic fowl (*Eimeria acervulina, Eimeria brunetti, Eimeria maxima, Eimeria mitis, Eimeria necatrix, Eimeria praecox*, and *Eimeria tenella*) was generated and also used as DNA template in PCR reactions with each of the turkey *Eimeria* species-specific primer pairs to determine if any of these *Eimeria* spp. of chickens could be amplified by the new primer pairs.

PCR primer design

The complete mtCOI CDS from five of the six *Eimeria* species listed above were extracted from complete mt genome sequences obtained for these parasites (Ogedengbe et al. 2014). A partial COI sequence for *E. innocua* was obtained from GenBank (1,257 bp, GenBank HG973049, see Vrba and Pakandl 2014). All available mtCOI sequences were aligned using a translation-based multiple alignment function within the bioinformatics package Geneious (V. 6.0 and later versions). The resulting alignment contained neither indels nor gaps; the complete mtCOI was 1,443 bp for the five *Eimeria* species for which complete CDS was available and translation products from all *Eimeria* species were well conserved. Setting each mtCOI sequence as a reference sequence sequentially permitted the identification of regions of low sequence identity that could be exploited for manual generation of primers.
Results

DNA samples Isolated DNA from each of the six Eimeria spp. infecting turkeys was successfully amplified using primers ITS-1 and ITS-2 of White et al. (1990) targeting the internal transcribed spacer 1 (ITS-1) region of the nuclear ribosomal RNA gene arrays. The amplifications of ITS region with ITS-1 and ITS-2 primer pair yielded a product of approximately 400 bp from each species. Primers Eim_COI_366F-M13F and Eim_COI_879R-M13R of El-Sherry et al. (2015) that target a portion of the mtCOI gene produced a PCR product of ~550 bp from all Eimeria species (data not shown).

PCR primer selection Six pairs of PCR primers designed to be species-specific with respect to the five Eimeria species of turkeys available for study were generated as follows: (1) E. adenooides-specific primers “E.ad.CO1_427F” (5′ CAAC CTCAGTAGATCTAATGTGA 3′) and “E.ad.CO1_1186R” (5′ GTGGAAGTGAGCAATGACA 3′) generate a 713-bp product; (2) E. dispersa-specific primers “E.disp.CO1_577F” (5′ A C A G C T T A T G T T A T G T G T 3’ 5′) and “E.disp.CO1_1028R” (5′ GCATACCAAGTATCTAATGAA 3′) generate a 861-bp product; (3) E. gallopavonis (861 bp), E. meleagrimitis (1,012 bp), and E. innocua with primer annealing temperatures of 62, 55, 62, 58, 52, and 50 °C, respectively (see Fig. 1 and Table 1). No inter-species cross-reactivity was observed at the optimized anneal temperatures; use of lower anneal temperatures or more permissive PCR reaction chemistries (e.g., higher divalent cation concentrations) may permit slight amplification of one or more non-target species (data not shown).

The amplifications of the COI region with species-specific primer pairs yielded the expected product sizes for E. adenooides (713 bp), E. meleagrimitis (554 bp), E. gallopavonis (861 bp), E. meleagridis (1,012 bp), E. dispersa (451 bp), and E. innocua (209 bp) as summarized in Table 1.

Table 1 Primer pairs, recommended PCR annealing temperatures, and expected PCR product sizes for the specific amplification of mitochondrial cytochrome c oxidase subunit I (COI) from 6 Eimeria species that infect turkeys

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Required anneal temp.</th>
<th>Sequence (5′ to 3′)</th>
<th>Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E.ad.CO1_427F</td>
<td>62 °C</td>
<td>5′-CCAAACCTCAGTAGATCTAATGTGA-3′</td>
<td>713</td>
</tr>
<tr>
<td>E.ad.CO1_1186R</td>
<td>55 °C</td>
<td>5′-GTGGAAGTGAGCAATGACA-3′</td>
<td>451</td>
</tr>
<tr>
<td>E.disp.CO1_577F</td>
<td>62 °C</td>
<td>5′-ACAGCTATATGTAATTGTGT-3′</td>
<td>861</td>
</tr>
<tr>
<td>E.disp.CO1_1028R</td>
<td>58 °C</td>
<td>5′- GCATACCAAGTATCTAATGAA-3′</td>
<td>554</td>
</tr>
<tr>
<td>E.gal.CO1_1153R</td>
<td>52 °C</td>
<td>5′- TTGAAGATTAGGGAAATATAA-3′</td>
<td>1,012</td>
</tr>
<tr>
<td>E.gal.CO1_431F</td>
<td>58 °C</td>
<td>5′- CTCAAGTATCTCCTCAG-3′</td>
<td>209</td>
</tr>
<tr>
<td>E.mel.CO1_1028R</td>
<td>50 °C</td>
<td>5′-TTCAAGATTAGGAAATATAA-3′</td>
<td>554</td>
</tr>
<tr>
<td>E.mel.CO1_474F</td>
<td>62 °C</td>
<td>5′-GCGTACCAGATATCTAAGG-3′</td>
<td>451</td>
</tr>
<tr>
<td>E.inn.CO1.396F</td>
<td>58 °C</td>
<td>5′-GCCATACCAAGTATCTAAGG-3′</td>
<td>1,012</td>
</tr>
<tr>
<td>E.inn.CO1.604R</td>
<td>52 °C</td>
<td>5′-GCCATACCAAGTATCTAAGG-3′</td>
<td>209</td>
</tr>
<tr>
<td>E.md.CO1_431F</td>
<td>62 °C</td>
<td>5′-TTCAAGATTAGGAAATATAA-3′</td>
<td>554</td>
</tr>
<tr>
<td>E.md.CO1_1443R</td>
<td>58 °C</td>
<td>5′-GCGTACCAGATATCTAAGG-3′</td>
<td>1,012</td>
</tr>
<tr>
<td>E.md.CO1_1153R</td>
<td>52 °C</td>
<td>5′-CTCAAGTATCTCCTCAG-3′</td>
<td>209</td>
</tr>
</tbody>
</table>

PCR optimizations The mtCOI species-specific primer sets successfully amplified single species template DNA of each of the six species at a wide range of annealing temperatures. The PCR reaction chemistries were optimized by adjusting magnesium chloride (MgCl2) and dNTP concentrations. Extension, annealing temperatures, and cycle parameters were optimized using combinations of mixed DNA samples containing all six Eimeria spp. infecting turkeys or only the five non-target Eimeria species as PCR templates (data not shown). Ultimately, target-specific PCR reactions were obtained for primer pairs specific for E. adenoideis, E. dispersa, E. gallopavonis, E. meleagrimitis, and E. innocua with primer annealing temperatures of 62, 55, 62, 58, 52, and 50 °C, respectively (see Fig. 1 and Table 1). No inter-species cross-reactivity was observed at the optimized anneal temperatures; use of lower anneal temperatures or more permissive PCR reaction chemistries (e.g., higher divalent cation concentrations) may permit slight amplification of one or more non-target species (data not shown).

The amplifications of the COI region with species-specific primer pairs yielded the expected product sizes for E. adenoideis (713 bp), E. meleagrimitis (554 bp), E. gallopavonis (861 bp), E. meleagridis (1,012 bp), E. dispersa (451 bp), and E. innocua (209 bp) as summarized in Table 1. The amplification bands of expected sizes were obtained only using the appropriate species-specific primer pair even when the target DNA (10 ng) was mixed with fivefold of mixed genomic DNA from the five other non-target species. No amplification was produced using any of the five primer pairs when 70 ng of mixed chicken Eimeria spp. (10 ng per Eimeria species) was used as the template DNA (Fig. 2).

Discussion

Molecular markers from the nuclear genome such as ribosomal ITS-1, ITS-2, and 18S rDNA have been tested previously...
for use in identifying Eimeria species infecting domestic turkeys and chickens (e.g., Stucki et al. 1993; Tsuji et al. 1997; Schnitzler et al. 1998; Woods et al. 2000; Gasser et al. 2001; Su et al. 2003; Rampin et al. 2006; Cook et al. 2010; Poplstein

Fig. 1 Aligned complete or near complete mitochondrial COI sequences from six Eimeria species infecting turkeys (from Ogedengbe et al. 2014 and Vrba and Pakandl 2014) with the locations of the various species-specific primers and resultant PCR products mapped onto each sequence.

Fig. 2 Specificity of species-specific PCR reactions targeting the mitochondrial cytochrome c oxidase subunit I (COI) loci of six Eimeria species that infect turkeys. Each species-specific primer pair was tested (using the anneal temperatures found in Table 1) for its ability to amplify the target Eimeria species in a mixed DNA template containing 10 ng total genomic DNA from each of E. meleagridis, E. gallopavonis, E. adenoeides, E. meleagrimitis, E. dispersa, and E. innocua (“Turkey Eimeria DNA”). Absence of amplification of non-target template DNA was tested using a mixture containing 10 ng total genomic DNA from the five Eimeria species that infect turkeys not target by the primer pair being tested (“Control DNA”). Additionally, each species-specific primer pair was also tested for absence of amplification when a mixture containing 10 ng total genomic DNA from each of seven Eimeria species that infect chickens (i.e., E. acervulina, E. brunetti, E. maxima, E. mitis, E. necatrix, E. praecox, and E. tenella; “Mixed DNA from seven Eimeria spp. from chickens”). Specific amplification of PCR products of the expected amplicon sizes (see Table 1) were only detected in PCR reactions that contained the target DNA.
and Vrba 2011; Ogedengbe et al. 2011; El-Sherry et al. 2013); likewise, RAPD/SCAR-based species-specific PCR assays have been described by Fernandez et al. (2003a, b) and Vrba and Pakandl (2014) for chickens and turkey Eimeria spp., respectively. The comparatively conserved nature of the nuclear 18 s rDNA locus has permitted evolutionary relationships among many taxonomic groups in the Apicomplexa, including chicken Eimeria spp., to be inferred successfully using 18 s rDNA sequences (e.g., Barta et al. 1997 and many others). In contrast, nuclear ribosomal ITS sequences varied distinctly among strains of a single species, as well as within a single parasite, due to polymorphisms among copies within the nuclear ribosomal gene arrays (e.g., Barta et al. 1998; Cook et al. 2010; Vrba and Pakandl 2014). The high rate of sequence divergence at the ITS loci makes the ITS regions suitable for work at the level of strains but this renders these loci as unreliable markers for use at the species level (Cruickshank 2002; Lew et al. 2003). Although nuclear 18S rRNA has clear utility for use at the genus level and higher taxonomic ranks, 18S rDNA is not highly suited for species identification within coccidia (El-Sherry et al. 2013); like the ITS loci, the nuclear 18S rDNA loci can have highly divergent paralogous copies within an individual coccidium that can confuse sequence-based genotyping of these parasites (e.g., Vrba et al. 2011; El-Sherry et al. 2013).

The mtCOI locus, in contrast, has been shown to be more effective as a species-level genetic marker for Eimeria spp. infecting turkeys and other galliform birds than nuclear 18S rDNA (Ogedengbe et al. 2011; Vrba et al. 2011; El-Sherry et al. 2013) due to comparatively faster genetic divergence within the former locus and lack of paralogous gene duplication, at least within the eimerid coccidia as far as is known (Ogedengbe et al. 2014). The current study describes a useful PCR-based diagnostic method for the detection of commonly documented species of Eimeria infecting turkeys. The described PCR targets overlapping regions of the mtCOI gene of each species. Despite the comparatively wide range of recommended anneal temperatures used for the primer pairs described in the present paper, it is possible to run PCR reactions using all six primer pairs in a single gradient thermocycler simultaneously.

The specificity of the primers developed in the present work was confirmed using a mixed species DNA template suggesting that these primer sets could be useful for conducting field surveys to describe the range of Eimeria species commonly encountered in the commercial poultry industry. Several attempts have been made to use genus-specific nuclear 18S rDNA or mtCOI primers in field samples containing multiple unidentified Eimeria species (e.g., Cook et al. 2010; Miska et al. 2010). In these cases, the resulting PCR products were cloned and sequenced as a means of estimating the species that were present in each sample. Unfortunately, a considerable portion of the resulting clones demonstrated chimeric sequences that apparently reflect PCR artefacts (see El-Sherry et al. 2013 for a more detailed explanation). RAPD/SCAR-based PCR assays for six Eimeria species infecting turkeys has been described recently (Vrba and Pakandl 2014) that depends on amplification and quantification of single copy, divergent loci within the nuclear genomes of these Eimeria species. A potential drawback of any RAPD/SCAR-based species-specific PCR assay is that they may be less sensitive than assays based on nuclear ribosomal DNA loci (e.g., 18S, ITS-1, ITS-2) that have many gene copies (likely more than 100 copies per nuclear genome in Eimeria species; see Logan-Klumper et al. 2012) within the nuclear genome (Vrba et al. 2010) or those assays, including the present work, based on mitochondrial loci. In the case of mitochondrial genes, these single copy genes are found in multiple mt genome copies in each life cycle stage of Eimeria species; for example, there are approximately 50-fold mt genome copies of E. tenella compared to the nuclear genome (Hikosaka et al. 2011).

In the case of RAPD-SCAR-based multiplex PCR for seven Eimeria species infecting chickens (Fernandez et al. 2003a, 2003b), the species-level stability of the sequence target is unknown for RAPD/SCAR-based assays (Vrba et al. 2010). Testing of the assay developed by Fernandez et al. (2003b) demonstrated that genomic DNA from a single species produced a species-specific product from DNA representing one or a few oocysts, but this sensitivity dropped to 20 or more oocysts when applied to field samples (Frölich et al. 2013). Further, Frölich et al. (2013) found that a mixed species template containing 1% of a contaminating species could not be detected by the assay when run as a multiplexed PCR reaction.

The PCR products generated with the COI species-specific primers described in the present work can be sequenced successfully to give DNA sequence that can be useful in a number of ways. Sequencing of any PCR product generated by this assay can provide data with utility for sequence-based genotyping as well as for phylogenetic analyses. This is not possible with the nu ITS loci because of copy to copy variations in template sequences, including widely divergent product lengths (e.g., Cook et al. 2010). The mtCOI locus is a protein-coding sequence without introns making amplification products of stable lengths that can be readily checked using translation to confirm that the sequenced product was of the correct genetic target.

The amplicons generated by the PCR assay described herein generate a minimum of 166 bp (in the case of the shortest product generated by the E. innocua-specific primer pair) found within a region of the mtCOI gene for which considerable comparative sequence information is available (amplified by primers 400 F/1202R, see El-Sherry et al. 2013). Thus, the PCR products can contribute to molecular phylogenetic analyses as well. In contrast, RAPD/SCAR-based assays (e.g., Fernandez et al. 2003a, b for Eimeria spp. of chickens or Vrba...
and Pakandi 2014 for Eimeria spp. of turkeys) do not amplify homologous loci thus sequence data cannot be compared among species. In the case of the present COI-based assay, a PCR failure with the species-specific primers can be backed up with genus-specific mtCOI primers (e.g., primers Eim_COI_366F-M13F and Eim_COI_879R-M13R of El-Sherry et al. 2014b) that amplify successfully a portion of the mtCOI gene from all turkey and chicken Eimeria species tested to date (see El-Sherry et al. 2015). Use of such genus-specific primers provides a useful positive control to establish the suitability of DNA templates being used for any of the species-specific PCR reactions and will permit characterization of Eimeria species of turkeys (e.g., E. subrotunda) and other hosts for which species-specific primers have yet to be developed. As the publically available dataset of partial and complete mtCOI sequences for various Eimeria species grows, the value of PCR-based identification, combined with sequence-based genotyping on the species-specific PCR amplicons from these parasites, will be enhanced.

Acknowledgments This work was supported by a Discovery Grant from the Natural Science and Engineering Council of Canada (NSERC) and an Applied Research and Commercialization (ARC) Grant from FedDev (Southern Ontario). In vivo work was supported, in part, by the Ontario Ministry of Food and Rural Affairs (OMAFRA). Special thanks to Dr. V. Vrbá (BIOPHARM, Institute of Biopharmacy and Veterinary Drugs, Pohorí-Chotouň, Jilove u Prahy 254 49, Czech Republic) for providing DNA from a European isolate of E. innocua.

References


Long PL (1973) Pathology and pathogenicity of coccidial infection. University Park Press, Baltimore, pp 251–294
