AUSTRALIAN POULTRY CRC

FINAL REPORT

Program Poultry Health / 2A Discovery and Development

Project No: Poultry CRC 09-27

PROJECT LEADER: Jess Morgan

DATE OF COMPLETION: 11/12/2009

Project 09-27

Eimeria species diagnostics based on non-nuclear genetic markers

ISBN 1 921010 47 9

Eimeria species diagnostics based on non-nuclear genetic markers
Project No. 09-27

Researcher Contact Details
Dr Jess A T Morgan
Department of Employment, Economic
Development and Innovation
Animal Research Institute
Locked Mail Bag #4
Moorooka Qld 4105
Phone: (07) 3362 9494
Fax: (07) 3362 9429
Email: Jessica.morgan@deedi.qld.gov.au

Australian Poultry CRC Contact Details
PO Box U242
University of New England
ARMIDALE NSW 2351

Phone: 02 6773 3767
Fax: 02 6773 3050
Email: poultrycrc@une.edu.au.
Website: http://www.poultrycrc.com.au

Published in March 2010

Executive Summary

This report describes the screening of two non-nuclear genomes for their use as molecular diagnostic markers to characterize species and strains of chicken *Eimeria*. The sequences obtained will provide alternative targets for assays to diagnose poultry coccidiosis. Improved diagnostic tools will benefit both meat and egg producers in the Australian poultry industry who suffer from coccidiosis outbreaks. They will also assist commercial live vaccine companies for both quality control and strain selection.

Coccidiosis in Australian flocks is caused by infection with one or more of the seven described species of *Eimeria*. The disease is costly for poultry industries around the world. Mixed species infections are common but overlapping morphological characters make species identification difficult. Molecular identification of species is more reliable but assays based on nuclear ribosomal RNA markers have been confounded by within strain variation. Non-nuclear genetic markers offer independent genomes that do not undergo recombination. As a result genes can acquire mutations more rapidly than nuclear DNA making them suitable for species and population genetic studies. There has been limited research to distinguish strains of *Eimeria* despite an urgent need within the industry to separate vaccine strains from wild strains.

This project was initiated to assess the utility of genes within the mitochondrial and apicoplast genomes to distinguish strains and species of *Eimeria*. A region of the RNA polymerase C2 gene (RpoC2) in the apicoplast genome and the cytochrome b gene of the mitochondrial genome were sequenced and compared for two strains of each of the seven species of *Eimeria*. The results of the study will determine if these genomes are suitable to develop diagnostic assays to benefit the Australian poultry industry by improving their ability to diagnose and monitor coccidiosis outbreaks in flocks.

Primers were designed from the published *E. tenella* apicoplast genome (Cai *et al.* 2003) and available sequences from the *Eimeria tenella* genome project (The Sanger Institute U.K.). DNA was extracted from pure lines of *Eimeria* species maintained at the Department of Employment, Economic Development and Innovation's (DEEDI) Animal Research Institute in Brisbane. Target DNA was amplified using polymerase chain reaction and directly sequenced. Sequences were edited and aligned and phylogenetic trees were constructed to investigate relationships among the species and strains.

The apicoplast RpoC2 gene could only be amplified in *E. tenella* and *E. necatrix* so two additional apicoplast genes were sequenced; the large subunit ribosomal RNA gene (LSU) and the translation elongation factor Tu gene (TufA). Sufficient variability was found in both the mitochondrial and apicoplast genes to distinguish the seven described *Eimeria* species of chickens. No within-strain variability was detected making these genes well suited to diagnostic assays. Low level between-strain divergence was detected in four species suggesting that the non-nuclear genomes may contain enough variability to investigate strain differences. Phylogenetic analyses separate the strains into well supported species groups and indicate *E. tenella* and *E. necatrix* are sister species that branched early in the lineage.

The implications of these findings are that the mitochondrial and apicoplast genomes will be suitable for developing diagnostic assays for the seven species of *Eimeria*. They may also contain enough mutations to distinguish among strains within a species. Better diagnostic markers would allow improved coccidiosis management, resulting in more efficient chicken meat and egg production, improved animal welfare through control of the disease and greater consumer confidence from improved product image.

Recommendations from this project are 1. develop new species diagnostic assays to target the mitochondrial genome, 2. determine the position of the new Operational Taxonomic Units (OTU's) described by Cantacessi *et al.*(2008) in the mitochondrial and apicoplast phylogenies and 3. sequence more of both genomes to develop diagnostic strain markers.

Contents

List of Tables

List of Figures

Introduction

Coccidiosis of chickens is caused by infection with species of *Eimeria*. Seven species of *Eimeria* have been recognised, based on parasite morphology, prepatent period, site of infection and pathogenicity (Callow 1984; Jorgensen *et al.* 1997) although new taxonomic units have been characterized recently (Cantacessi *et al.* 2008). With the advent of live *Eimeria* vaccines specific diagnosis of infection has a key role in the prevention, surveillance and control of coccidiosis (Morris & Gasser 2006). Species-level diagnosis based on morphological characters can be unreliable because diagnostic characters such as oocyst size, pre-patent period and minimum sporulation times overlap among some *Eimeria* species, as do the sites of infection within the intestine (Long & Joyner 1984). Thus, molecular tools are increasingly being developed and relied upon for the species-level diagnosis of *Eimeria*. Pioneering work using conventional polymerase chain reaction (PCR) with nucleotide sequence-derived markers has greatly improved specific diagnosis (Procunier *et al.* 1993; Barta *et al.* 1997; Schnitzler *et al.* 1998; Schnitzler *et al.* 1999; Fernandez *et al.* 2003; Lew *et al.* 2003; Gasser *et al.* 2005; Lien *et al.* 2007). Although clear pathogenic differences exist among strains of *Eimeria* there is an urgent need to develop genetic markers to distinguish them. These markers could be used to track drug resistant strains and separate vaccine and field strains to better understand outbreaks.

Different gene regions of *Eimeria* offer a range of divergence levels for comparing species and strains. The nuclear ribosomal RNA 18S (SSU) gene is highly conserved and is useful for investigating evolutionary relationships at the genus level such as host-parasite co-evolutionary patterns (Barta *et al.* 1991). In contrast the nuclear ribosomal internal transcribed spacers are so variable that sequences obtained from a single strain do not necessarily group together. These hypervariable genes can still be useful for distinguishing species but their evolutionary histories are difficult to infer (Lew *et al.* 2003; Blake *et al.* 2006; Swinkels *et al.* 2006; Cantacessi *et al.* 2008; Morgan *et al.* 2009; Schwarz *et al.* 2009). Diagnostic assays (at strain or species level) are easier to develop and interpret if they are based on single nucleotide polymorphisms (SNPs) as opposed to allele frequency differences. Multicopy genes such as those found in ribosomal DNA, mitochondrial DNA or apicoplast DNA are preferred over single copy nuclear genes because they amplify more readily and tend to be conserved within an individual (i.e. an individual caries multiple copies of only one allele due to either concerted evolution or maternal inheritance – unfortunately the ribosomal DNA of poultry *Eimeria* does not follow a concerted evolution model). In a diploid organism, an individual can carry two different alleles per nuclear gene greatly complicating assay interpretation. Selecting the appropriate molecular marker or gene region for resolving a particular systematic question is a difficult process. Ribosomal DNA is frequently used for species level comparisons while mitochondrial DNA is commonly used to detect strain or population differences. Microsatellites are more variable again so are used for population comparisons but they are more complicated to analyse than SNPs. As the *E. tenella* genome sequencing project progresses and bioinformatics platforms advance it will be possible to mine the complete genome for strain or species diagnostic SNPs. The mitochondrial and apicoplast genomes may provide an intermediate level of diversity. A recent study by Schwarz *et al*. (2009) of the mitochondrial cytochrome oxidase 1 (CO1) gene found that this marker separated strains of *E. maxima*, *E. tenella* and *E. acervulina* by species. They also reported some intra-specific diversity within each of the species.

Polymerase Chain Reaction (PCR) methodology for DNA amplification and direct automated DNA sequencing are efficient and rapid techniques commonly used in species diagnostics and population genetics. Genetic markers present in multiple copies within a genome amplify more easily than single copy genes but the copies must be conserved for direct sequencing otherwise an additional cloning step is necessary adding time and cost to screening. Unfortunately not all of the ribosomal RNA copies (roughly 140 copies per genome based on the *Eimeria tenella* genome project) evolve in concert making direct sequencing of this marker difficult (Cantacessi *et al.* 2008).

In addition to the nuclear genome *Eimeria* cells carry two non-nuclear genomes in their cytoplasm, one in their mitochondria and one in their apicoplast. Both genomes are multi-copy within an individual and both are maternally inherited (i.e. they do not undergo recombination but are passed directly from mother to daughter cells). Without recombination these genomes can acquire mutations more rapidly than the nuclear genome making them better suited to comparisons at a species and strain level (Crozier 1990; Oborník *et al.* 2002). The sequencing of the *Eimeria tenella* (Houghton strain) genome currently in process at the Sanger Institute UK (Shirley *et al.* 2004) and the publication of the complete apicoplast genome of *E. tenella* (Penn State) (Cai *et al.* 2003) has facilitated investigation of these non-nuclear genomes for the development of diagnostic species and strain level markers. In *E. tenella* (Houghton) the complete mitochondrial genome is 6.2 kilo bases (Sanger Institute UK) and in *E. tenella* (Penn State) the complete apicoplast genome is 34.7 kilo bases (Cai *et al.* 2003).

The aim of this short project was to assess the suitability of two non-nuclear genetic markers for distinguishing among species and strains of *Eimeria* that infect chickens. The cytochrome b gene (Cytb) within the mitochondrial genome and the RNA polymerase C2 gene (RpoC2) within the apicoplast genome were targeted.

Objectives

Sequence two 800 base pair regions, one in the apicoplast genome (RpoC2 gene) and one in the mitochondrial genome (cytochrome b), for two strains of each of the seven species of chicken *Eimeria*.

The objective of this project was to improve management of coccidiosis in poultry by identifying if non-nuclear genetic markers could be used to identify and differentiate strains and species of *Eimeria* that infect chickens. To achieve this objective, a region of the apicoplast genome and a region of the mitochondrial genome were sequenced and compared. The ability to distinguish among strains and species of *Eimeria* will assist the poultry industry to track the occurrence and spread of coccidiosis and better monitor control measures.

Methodology

Samples and DNA extraction

Queensland Primary Industries and Fisheries, within the Department of Employment, Economic Development and Innovation maintains pure Australian strains of each of the seven *Eimeria* species as purified oocysts stored in 2% w/v potassium dichromate. Prior to DNA extraction 10^6 oocysts were washed with distilled water then they were resuspended in 80 µL of phosphate-buffered saline (PBS, pH 7.2) and their walls were disrupted by homogenization with 0.1 g of 1 mm glass beads for 5 min in a MiniBeadbeater-96 (Biospec Products, Bartlesville, OK, USA). DNA was extracted from the cracked oocysts using a DNeasy Tissue Kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions. Details of the species and strains used in this study are provided in Table 1.

Table 1 Species and strains of *Eimeria* included in marker study

Species	Strain	Origin
E. tenella	Redten	Queensland
	Medten	Victoria
	Ingten	NSW
	Penn State-ten*	USA
	Houghton-ten*	United Kingdom
E. necatrix	Mednec	Victoria
	Gronec	NSW
E. acervulina	Royace	Queensland
	Ponace	Queensland
	Unknown-ace*	Unknown likely USA
E. maxima	Medmax	Victoria
	Ingmax	Victoria
	Unkown-max*	Unknown likely USA
E. brunetti	Bowbru	NSW
	Andbru	Queensland
	Monbru	South Australia
E. praecox	Jorpra	Queensland
	Ingpra	NSW
E. mitis	Jormit	Queensland
	Redmit	Queensland
*Eimeria falciformis**	Unknown*	Unknown

* Publically available sequences

Primer design, DNA amplification and sequencing

Complete mitochondrial and partial apicoplast genome sequences of the Houghton Strain of *Eimeria tenella* are publically available from The Sanger Institute (UK). The complete apicoplast genome has also been sequenced for a Penn State Strain of *E. tenella* (Cai *et al.* 2003). Primers were either designed from the *E. tenella* genome or they were based on those published in Cai *et al.*(2003).

Primers targeted the mitochondrial cytochrome b gene (Cytb) and three apicoplast genes, the RNA polymerase C2 gene (RpoC2), the large subunit RNA gene (LSU) and the translation elongation factor Tu gene (TufA).

Amplification reactions for all genes were carried out in 10 µl volumes containing 0.5 µM of each primer pair, combined with 10-100 ng of extracted DNA, 10x HotMaster Taq buffer (Eppendorf, Australia, containing 25 mM magnesium), 0.8 mM dNTP, and 0.05 units/µl of HotMaster Taq DNA polymerase (Eppendorf, Australia). Thermal cycling conditions consisted of an initial denaturation (95°C for 2 minutes) followed by 30 cycles of 95°C for 30 seconds, annealing at 53°C (Cyt b and TufA) or 52°C (LSU and RpoC2) for 30 seconds and extension at 72°C for 1 minute 30 seconds, with a final extension step of 72°C for 7 minutes. Cycling was performed in a DNA Engine Gradient Cycler (Bio-Rad Laboratories, Gladesville, New South Wales). PCR products were viewed on a 2% agarose TBE gel stained with GelRed (Biotium, USA). PCR products were desalted prior to sequencing using

Exosap-it® (USB Corporation distributed by GE Healthcare Bio-Sciences, Rydalmere NSW, Australia). Approximately 20ng of DNA was used in standard ABI Dye Terminator sequencing reactions using Big Dye Vers 3.1 technology (Applied Biosystems, California) and were run on an Applied Biosystems 3130xl Genetic Analyser (Griffith University DNA Sequencing Facility, School of Biomolecular and Biomedical Science, Griffith University, Qld, Australia). Forward and reverse sequences were edited and aligned using Sequencher (Vers 4.7 Gene Codes Corporation, Ann Arbor, MI, USA).

Phylogenetic analyses

Sequences for each gene were aligned using ClustalX (Vers 1.81, Thompson *et al.* 1997) then alignments were eyeball edited prior to export in nexus format for phylogenetic analyses. Pairwise distances using uncorrected p-distances were calculated in PAUP* (Vers 4.0b10, Swofford 2001). Phylogenetic trees were constructed using maximum parsimony (P), maximum likelihood (L) and distance matrix analyses (D) in PAUP* (Vers 4.0b10, Swofford 2001). Before generating the trees a series of likelihood ratio tests were completed using MrModeltest (Vers 2.3, Nylander 2004) to determine the best nucleotide substitution model to use for likelihood and distance analyses. For the mtDNA cytochrome b gene a general time reversible model (GTR) with among site heterogeneity (G) was selected (summarized as GTR+G) with the following command block for use in Paup*.

MtDNA Cytochrome b (Cytb)
BEGIN PAUP;
Lset Base=(0.2915 0.1663 0.1461) Nst=6 Rmat=(3.2223 5.4141 2.9782 0.9142 14.5131) Rates=gamma Shape=0.2491 Pinvar=0;
END;

A general time reversible model with among site heterogeneity (GTR+G) was also selected for the apicoplast large subunit RNA (LSU) and the translation elongation factor Tu gene (TufA) with the following command blocks for use in Paup*.

Apicoplast Large subunit RNA (LSU)
BEGIN PAUP;
Lset Base=(0.4120 0.1211 0.1452) Nst=6 Rmat=(1.3935 3.7023 3.3372 0.0807 6.4020) Rates=gamma Shape=0.5301 Pinvar=0;
END;

Apicoplast translation elongation factor Tu gene (TufA) gene
BEGIN PAUP;
Lset Base=(0.4121 0.0865 0.1563) Nst=6 Rmat=(5.2195 7.3748 3.1869 3.9801 9.9860) Rates=gamma Shape=0.3942 Pinvar=0;
END;

Unweighted trees were found using heuristic searches with random sequence addition and tree-bisection-reconnection (TBR) branch swapping. For the parsimony analysis gaps were treated as missing data. Other settings used were Mulpars in effect, Maxtrees set to 1000 (P) or 200 (D and L) and heuristic search repetitions were set to 1000 (P) or one (D and L). Support for nodes was assessed using bootstrap resampling (P 1,000 replicates, D 1,000 replicates, L 100 replicates due to high computational demands). Where publically available sequences were obtainable (Cytb and LSU) *Eimeria falciformis*, an intestinal protozoan of rats and mice, was included as outgroup to root the trees.

Results

Primer design, DNA amplification and sequencing

Mitochondrial and apicoplast primer sequences and positions are shown in Table 2. Additional genes were screened for the apicoplast genome after experiencing repeated failures trying to amplify a product for the RpoC2 gene. The additional genes screened were amplified using primers from (Cai *et al.* 2003) and included the translation elongation factor Tu gene (TufA) and the large subunit RNA (LSU) gene (Table 2).

Table 2 Primers and their targets amplified in the marker study. Position in the mitochondrial genome corresponds to the 5' position in the Sanger genome contig 00018452. Position in the apicoplast genome represents the 5' sequence position in genbank accession AY217738.

Name	Sequence 5' to 3'	Target gene	Origin	Position
Mitochondrial genome				
Cytb F	GTTTATTATGTCTCAAGTGAGATC	Cytb	Authors	6191
Cytb R	ATACCTAATTCTTTATGGTTTGC	Cytb	Authors	7356
Apicoplast genome				
RpoC2 F1	TATATTTTAATAAAATGTCTATGCC	RpoC2	Authors	19882
RpoC2 R1	TAATTCCTAAAAATATAGTATCGC	RpoC2	Authors	20930
RpoC2 F2	TTATTATTTATAAATTGTCTGTGT	RpoC2	Authors	20876
RpoC2 R2	AAGAACTTTTCATACAGGTGGAAC	RpoC2	Authors	22058
RpoC2 F3	CTTTTTATTTTATGGTATTTAACAC	RpoC2	Authors	21921
RpoC2 R3	TTACTTTTTTTTCTACAGGACCT	RpoC2	Authors	23031
RpoC2 F4	AATTTCTATAAGAATTAATCCTG	RpoC2	Authors	20644
RpoC2 R4	CTATATTATTTAATAAATATTTTCCG	RpoC2	Authors	21591
RpoC2 F5	CAGGAATAATAATATTATTAGGTG	RpoC2	Authors	20406
RpoC2 R5	GAGGTTATTTAGCAAATGCTACA	RpoC2	Authors	22635
RpoC2 F6	TATATTTTTATATAAAAATTCTCCAG	RpoC2	Authors	21650
RpoC2 R6	ATTCTATTAAAGATATTACATCAGG	RpoC2	Authors	20687
RpoC2 F7	TAATATTCCTTTTCTTCCACCG	RpoC2	Authors	22514
TufA F	GCTATTTTAGTTGTTTCTGCTAC	TufA	Cai *et al*. 2003	13785
TufA R	CATTATCTCCTGCTTCTGCCG	TufA	Cai *et al*. 2003	14322
LSU F	TTTGATAAACAGTCGCTTGG	LSU	Cai *et al*. 2003	3648
LSU R	AAATAGAAGTGAAAATGTCAGC	LSU	Cai *et al*. 2003	4094

Sequence alignments and comparison of genes

Most of the primers designed from the *E. tenella* (Penn State) RpoC2 gene failed to amplify any product at all, even when PCR conditions were relaxed (annealing temperature lowered to 45°C and an additional 10 cycles added). From every possible primer combination tested only two primer pairs amplified products in the RpoC2 gene. These primers only amplified products for strains of *E. tenella* and *E. necatrix*; bands could not be amplified for any of the other species. The first fragment was a 233 base pair product between primers RpoC2 F5 and RpoC2 R6 and the second fragment was a 473 base pair product between primers RpoC2 F7 and RpoC2 R3. *Eimeria tenella* and *E. necatrix* differed from each other through both fragments (6 single nucleotide polymorphisms (SNPs) in the F6-R6 fragment and 15 SNPS in the F7-R3 fragment). No strain differences were detected for either species, in either fragment. Due to the lack of success amplifying a product in the RpoC2 gene a

further two apicoplast genes, the large subunit RNA (LSU) and translation elongation factor Tu gene (TufA) genes were screened.

Comparing the different genes in a divergence table (Table 3) shows that at a strain level the genes were very conserved. Some divergence was seen among strains of *E. tenella* (Cytb and TufA), *E. acervulina* (LSU), *E mitis* (TufA) and *E. brunetti* (LSU). Across all of the genes the closest species (lowest percent divergence) were *E. tenella* and *E. necatrix*. The remaining five species diverged from each other across all genes by approximately 9% with none showing consistent close affiliations. The most variable gene was the apicoplast LSU although the estimate of 18.5% between the *E. tenella* and *E. necatrix* clade to the 5 species clade is probably an overestimate due to ambiguous bases from some poor sequence reads. Divergence in the RpoC2 gene fell within the range of that observed for the other genes between *E. tenella* and *E. necatrix*.

Table 3 Average percent divergence (using untransformed p-distances) within and between species of *Eimeria* for mitochondrial and apicoplast genes. Some comparisons were not possible (na).

	mean % divergence			
Genome	mtDNA	Apicoplast	Apicoplast	Apicoplast
gene	Cyt b	LSU	TufA	RpoC2
size of product (base pairs)	1123	347	479	706ψ
Within *Eimeria* species				
E. tenella	0.4	0	0.2	0
E. necatrix	0	na	na	0
E. maxima	0	na	na	na
E. acervulina	0	0.6	na	na
E. mitis	0	0	0.6	na
E. brunetti	0	2.0*	0	na
E. praecox	0	0	na	na
Between *Eimeria* species				
E. tenella to *E. necatrix*	1.7	6.1*	3.3	3.0
Within 5 species clade	9.9	9.1*	9.4	na
E. tenella & *E. necatrix* clade to 5 species clade	11.0	18.5*	10.6	na
To outgroup				
Eimeria to *E. falciformis*	11.0	19.7*	na	na

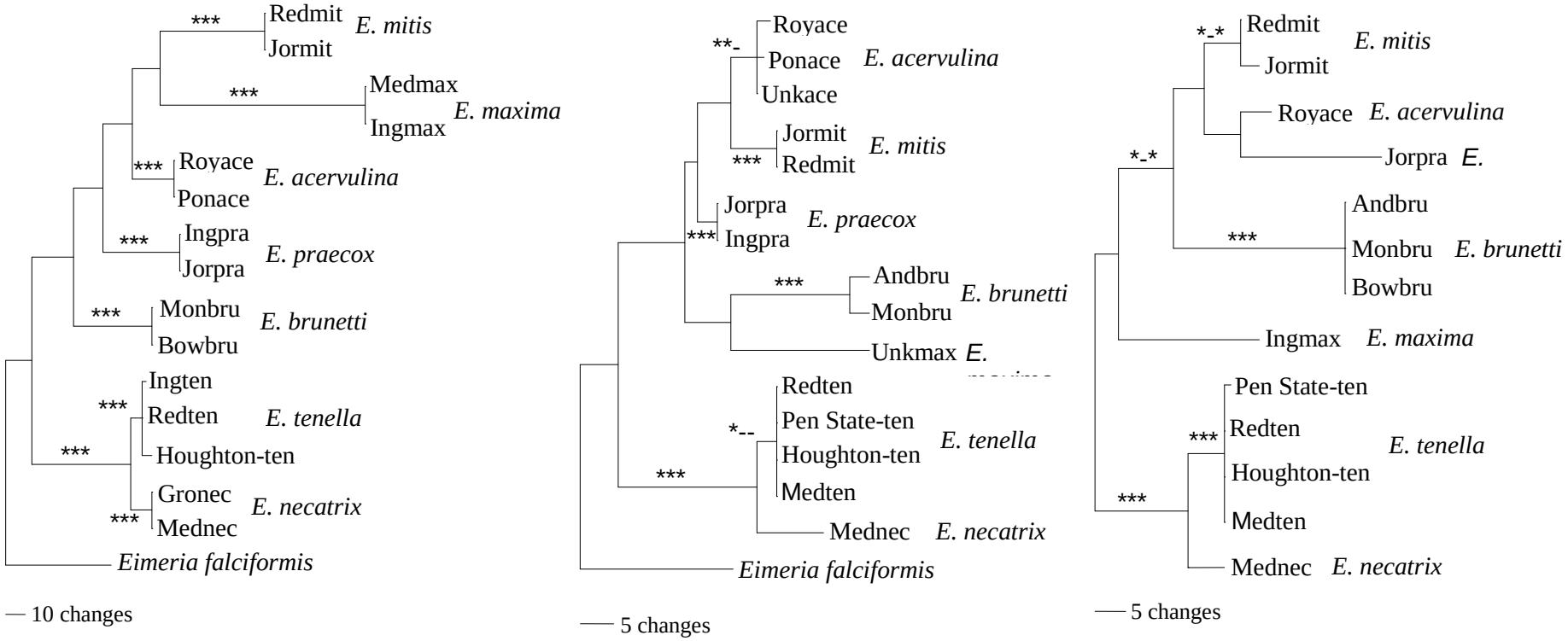
ψ 2 fragments combined
*ambiguous bases in sequences likely inflating scores

Phylogenetic trees

A 1123 base pair mitochondrial DNA cytochrome b product was amplified for two strains of each of the seven species of *Eimeria*. Phylogenetic analyses produced one parsimony tree of 458 steps (Figure 1a). The alignment contained 275 parsimony informative characters and the tree had a rescaled consistency index of 0.6420. For the distance analyses a single optimal tree was found with minimum evolution score 0.62414. Likelihood analyses also found a single tree with –Ln likelihood score 3597.38. All three tree building methods grouped the strains by species and also separated *E. tenella* and *E. necatrix* from the remaining species. Relationships among species within the 5 species clade were poorly resolved as reflected by the lack of bootstrap support for deeper nodes.

A 347 base pair region of the apicoplast LSU gene was obtained for at least one strain of each of the seven species (Australian *E. maxima* strains failed to produce a clean product for direct sequencing but sequence from an unknown (likely USA) strain was downloaded from Genbank). Only partial sequences were obtained from *Eimeria necatrix* (Mednec, 233 base pairs) and *Eimeria brunetti* (Monbru, 328 base pairs). The apicoplast LSU gene alignment contained 63 parsimony informative characters. Thirteen equally parsimonious trees were found of 149 steps and rescaled consistency index 0.7494 (one parsimony tree shown in Figure 1b). Six optimal distance trees were identified with minimum evolution score 0.56862 and three equally likely trees were found with –Ln likelihood score 1152.09407. As with the mitochondrial gene, all analyses of the apicoplast LSU gene grouped the *Eimeria* strains by species. Good bootstrap support was also observed separating *E. tenella* and *E. ecatrix* from the remaining species. Most of the variability observed among the optimal trees and tree building methods occurred in the 5 species clade containing *E. acervulina, E. mitis, E. praecox, E. brunetti* and *E. maxima*.

A 479 base pair region of the TufA gene was amplified for at least one strain of each of the seven species. No *Eimeria falciformis* TufA sequence was available for use as outgroup so tree rooting was forced (placing *E. tenella* and *E. necatrix* sister to the remaining species) following the same placement as seen in the LSU and Cytb gene tees. The TufA alignment contained 75 parsimony informative characters and two equally informative parsimony trees were found of 184 steps and rescaled consistency index of 0.7362 (one parsimony tree shown in Figure 1c). Analysis using minimum evolution identified six trees of score 0.51131. Likelihood analysis retained a single tree with –Ln likelihood score 1486.67584. Again *E. tenella* and *E. necatrix* fell as well supported sister taxa to the five other *Eimeria* species. Strains also separated by species and again most of the tree variability fell within the 5 species clade.



a. Mitochondrial DNA Cytb b. Apicoplast DNA LSU c. Apicoplast DNA Tuf A

Figure 1 Parsimony trees for one mitochondrial (a) and two apicoplast genes (b and c). Bootstrap support of greater than 85% is indicated above branches with a * for parsimony followed by distance then likelihood analyses.

Discussion

The *Eimeria* apicoplast genome is very AT rich (79%) which made primer design more difficult (avoiding mis-priming and primer dimers) than the mitochondrial genome (64% AT). Many attempts were made to design RpoC2 primers but none could be designed to amplify a product from all seven species. Surprisingly the majority of primers failed to amplify a product from *E. tenella*, the species whose sequence they were designed from. This result would suggest enormous strain and species differences but this was not found to be the case when multiple sequences were aligned. Two small products within the RpoC2 gene were amplified from *E. tenella* and *E. necatrix;* no strain differences were detected and the two species diverged by only 3%. Given this level of diversity it is unclear why amplification of the gene was so problematic. Within the apicoplast the RpoC2 gene is extremely AT rich (87%). This nucleotide bias may be creating unusually strong secondary structural configurations which could be disrupting the polymerase replication process. Cai *et al*. (2003) reported a low amino acid identity (39%) for RpoC2 against *Toxoplasma gondii* suggesting that the gene may still be worth investigating for species and strain differences, perhaps in a stepwise fashion from more conserved surrounding genes. The two alternative apicoplast genes targeted, LSU and TufA amplified more consistently, despite showing greater divergence for the *E. tenella* and *E. necatrix* comparison, than the partial RpoC2 gene.

Both the mitochondrial and apicoplast genomes provided sufficient sequence diversity to readily distinguish the seven species of *Eimeria*. In the mitochondrial Cytb gene species diverged from 1.7 to 11%. Through the apicoplast TufA gene, species divergence ranged from 3.3 to 10.6% and in the apicoplast LSU, species diversity ranged from 6.1 to 18.5%. The diversity observed in the LSU is an overestimate due to ambiguous positions within some of the sequences; it is likely that with cleaner sequences the range will fall below that seen in TufA (based on amino acid identity comparisons between *E. tenella* and *Toxoplasma gondii* genes made by (Cai *et al*. 2003)). Intraspecific strain differences were rare, mostly below 0.6%.

Phylogenetic analyses of the three gene regions support a close relationship between *E. tenella* and *E. necatrix*. This result is consistent with published nuclear ribosomal RNA 18S, ITS1 and ITS2 trees (Barta *et al.* 1997; Lew *et al.* 2003; Schwarz *et al.* 2009). The Cytb and LSU genes also support the basal branching position of *E. tenella* and *E. necatrix* within the *Eimeria* clade (no outgroup for TufA to determine the root of the tree). This finding is in agreement with the nuclear ribosomal RNA 18S phylogeny (Barta *et al.* 1997) and provides further evidence that trees based on the nuclear ribosomal RNA internal transcribed spacers are confounded by outgroup selection (Lew *et al.* 2003; Schwarz *et al.* 2009). The position of *E. maxima* within the five species clade is the least resolved. Long branch attraction (when the same mutation occurs at the same site by chance) rather than evolutionary relatedness may be responsible for its shifting position within the clade. Poor bootstrap support below species level in the five species clade prevents any firm interpretation on species relatedness. The nuclear markers also differ in their placement of these species with respect to one another within the genus. Phylogenetic trees force taxa into a bifurcating topology; it is possible that the radiation of the species in this clade was triggered by a single, or several closely timed events such that the clade would be better represented with a star topology.

Several applications of mitochondrial and apicoplast genome sequences can be investigated based on the results of this pilot study. Although present in multiple copies, the genes appear to be conserved within a strain making it possible to PCR amplify and direct sequence a single product per strain (not the case with the nuclear ribosomal RNA internal transcribed spacers). This makes sequencing samples fast and relatively cheap. Both the mitochondrial and apicoplast genomes contain sufficient variation to develop diagnostic assays for *Eimeria* species separation. Species show at least three times the level of divergence as strains, so matching an unknown strain to a species group should be straight forward. An interesting comparison to make would be to investigate where the three new operational taxonomic units (OTU's), X, Y and Z reported by Cantacessi *et al.*(2008) fall with respect to the described species. Phylogenetic placement of these OTU's based on nuclear ribosomal RNA internal

transcribed spacer 2 sequences is uncertain. It is possible that the mitochondrial and apicoplast genomes contain sufficient diversity to distinguish between strains of *Eimeria*. Unique mitochondrial Cyt b gene sequences were obtained for three strains of *E. tenella*. Similarly the apicoplast TufA gene separated the two Australian *E. tenella* strains from the Houghton and Penn State strains. Strain divergence is sufficiently rare that it is unlikely that a single gene will separate all strains for all species, however, recent work by Schwarz *et al.* (2009) comparing the mitochondrial CO1 gene from a large number samples found that strains of *E. maxima* varied up to 5.9%, strains of *E. acervulina* varied up to 1.4% and strains of *E. tenella* diverged up to 1.4%. This is a promising result for three species and suggests that the non-nuclear genomes are worth pursuing further for strain characterisation.

Implications

This project has found sufficient variability in two non-nuclear genetic markers to distinguish the seven described *Eimeria* species of chickens. No within-strain variability was detected in either genome making these genes better suited to diagnostic assays than the ribosomal RNA internal transcribed spacers currently used. Low level between-strain divergence was detected in four species suggesting that the non-nuclear genomes may contain enough variability for strain diagnostics. The ability to distinguish strains will permit a better understanding of the diversity of Australian *Eimeria* and these markers provide a tool for measuring a baseline on which to map virulence characters and outbreak patterns. This in turn could identify new strains to target as vaccine candidates, or assist in the selection of challenge strains.

Recommendations

The level of variability found in the mitochondrial and apicoplast genes makes them promising candidates for future research.

- Investigate where the new OTU's described by Cantacessi *et al.*(2008) fall with respect to the seven described *Eimeria* species.
- Develop new species diagnostic assays to target the mitochondrial genome.
- Sequence more of both genomes to find and develop diagnostic strain markers.

This research is designed to provide Industry with a range of diagnostic tests of different resolving power. Increasing resolution generally correlates with increasing cost. Selecting the appropriate diagnostic test will come down to the question being asked. Basic detection of *Eimeria* infection requires a microscope, and basic training, and is cheap. Highly sensitive assays that can diagnose individual species and quantify oocyst load directly from faecal samples with mixed-species infections (RT-PCR) requires specialized equipment, reagents and training so is more expensive. If industry is interested in increasing their diagnostic toolkit's resolving power to distinguish a wild outbreak strain of *Eimeria* from a vaccine strain then it should proceed with screening new markers and developing new assays.

Acknowledgments

I would like to thank Bartosz Wlodek for his technical assistance in the laboratory and Anthea Bruyeres for providing administrative support.

References

Barta, JR, Jenkins, MC & Danforth, HD 1991, 'Evolutionary relationships of Avian *Eimeria* species among other Apicomplexan Protozoa: Monophyly of the Apicomplexa is supported', *Molecular Biology and Evolution*, vol. 8, pp. 345-355.

Barta, JR, Martin, DS, Liberator, PA, Dashkevicz, M, Anderson, JW, Feighner, SD, *et al.* 1997, 'Phylogenetic relationships among eight *Eimeria* species infecting domestic fowl inferred using complete small subunit ribosomal DNA sequences', *Journal of Parasitology*, vol. 83, pp. 262-271.

Blake, DP, Hesketh, P, Archer, A, Shirley, MW & Smith, AL 2006, '*Eimeria maxima*: The influence of host genotype on parasite reproduction as revealed by quantitative real-time PCR', *International Journal for Parasitology*, vol. 36, pp. 97-105.

Cai, XM, Fuller, AL, McDougald, LR & Zhu, G 2003, 'Apicoplast genome of the coccidian *Eimeria tenella*', *Gene*, vol. 321, pp. 39-46.

Callow, LL 1984, *Animal Health in Australia: Volume 5 Protozoal and rickettsial diseases.*, Australian Government Publishing Service, Canberra.

Cantacessi, C, Riddell, S, Morris, GM, Doran, T, Woods, WG, Otranto, D, *et al.* 2008, 'Genetic characterization of three unique operational taxonomic units of *Eimeria* from chickens in Australia based on nuclear spacer ribosomal DNA', *Veterinary Parasitology*, vol. 152, pp. 226-234.

Crozier, RH 1990, 'From population genetics to phylogeny: uses and limits of mitochondrial DNA', *Australian Systematic Botany*, vol. 3, pp. 111-124.

Fernandez, S, Pagotto, AH, Furtado, MM, Katsuyama, M, Maderia, AMBN & Gruber, A 2003, 'A multiplex PCR assay for the simultaneous detection and discrimination of the seven *Eimeria* species that infect domestic fowl', *Parasitology*, vol. 127, pp. 317-325.

Gasser, RB, Skinner, R, Fadavi, R, Richards, G & Morris, G 2005, 'High-throughput capillary electrophoresis for the identification and differentiation of seven species of *Eimeria* from chickens', *Electrophoresis*, vol. 26, pp. 3479-3485.

Jorgensen, WK, Stewart, NP, Jeston, PJ, Molloy, JB, Blight, GW & Dalgliesh, RJ 1997, 'Isolation and pathogenicity of Australian strains of *Eimeria praecox* and *Eimeria mitis*', *Australian Veterinary Journal*, vol. 75, pp. 592-595.

Lew, AE, Anderson, GR, Minchin, CM, Jeston, PJ & Jorgensen, WK 2003, 'Inter- and intra-strain variation and PCR detection of the internal transcribed spacer 1 (ITS-1) sequences of Australian isolates of *Eimeria* species from chickens', *Veterinary Parasitology*, vol. 112, pp. 33-50.

Lien, YY, Sheu, SC, Liu, HJ, Chen, SC, Tsai, MY, Luo, SC, *et al.* 2007, 'Cloning and nucleotide sequencing of the second internal transcribed spacer of ribosomal DNA for three species of *Eimeria* from chickens in Taiwan', *The Veterinary Journal*, vol. 173, pp. 184-189.

Long, PL & Joyner, LP 1984, 'Problems in the Identification of Species of *Eimeria*', *Journal of Protozoology*, vol. 31, pp. 535-541.

Morgan, JAT, Constantinoiu, CC & Jorgensen, WK 2009, 'New diagnostic assays to improve control of coccidiosis in poultry', Rural Industries Research and Development Corporation, Canberra.

Morris, GM & Gasser, RB 2006, 'Biotechnological advances in the diagnosis of avian coccidiosis and the analysis of genetic variation in Eimeria', *Biotechnology Advances* vol. 24, pp. 590-603.

Nylander, JAA 2004, MrModeltest 2.3. Program distributed by the author Evolutionary Biology Centre, Uppsala University.

Oborník, M, Van de Peer, Y, Hypsa, V, Frickey, T, Slapeta, JR, Meyer, A, *et al.* 2002, 'Phylogenetic analyses suggest lateral gene transfer from the mitochondrion to the apicoplast', *Gene*, vol. 285, pp. 109-118.

Procunier, JD, Fernando, MA & Barta, JR 1993, 'Species and strain differentiation of *Eimeria* spp. of the domestic fowl using DNA polymorphisms amplified by arbitrary primers', *Parasitological Research*, vol. 79, pp. 98-102.

Schnitzler, BE, Thebo, PL, Mattsson, JG, Tomley, FM & Shirley, MW 1998, 'Development of a diagnostic PCR assay for the detection and discrimination of four pathogenic *Eimeria* species of the chicken', *Avian Pathology*, vol. 27, pp. 490-497.

Schnitzler, BE, Thebo, PL, Tomley Fiona M., Uggla, A & Shirley, MW 1999, 'PCR Identification of chicken *Eimeria*: a simplified read-out', *Avian Pathology*, vol. 28, pp. 89-93.

Schwarz, RS, Jenkins, MC, Klopp, S & Miska, KB 2009, 'Genomic analysis of *Eimeria spp*. populations in relation to performance levels of broiler chicken farms in Arkansas and North Carolina', *The Journal of Parasitology*, vol. 95, pp. 871-880.

Shirley, M, Ivens, A, Gruber, A, Madeira, AM, Wan, KL, Dear, PH, *et al.* 2004, 'The *Eimeria* genome projects: a sequence of events', *Trends in Parasitology*, vol. 20, pp. 199–201.

Swinkels, WJC, Post, J, Cornelissen, JB, Engel, B, Boersma, WJA & Rebel, JMJ 2006, 'Immune responses in *Eimeria acervulina* infected one-day-old broilers compared to amount of *Eimeria* in the duodenum, measured by real-time PCR', *Veterinary Parasitology*, vol. 138, pp. 223-233.

Swofford, DL 2001, PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Associates, Sunderland, Massachusetts.

Thompson, JD, Gibson, TJ, Plewniak, F, Jeanmougin, F & Higgins, DG 1997, 'The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools', *Nucleic Acids Research*, vol. 25, pp. 4876-4882.

Plain English Compendium Summary

Project Title:	
Project No.:	Poultry CRC 09-27
Researcher:	Dr Jess Morgan
Organisation:	Department of Employment, Economic Development and Innovation Agri-Science Queensland Animal Research Institute Locked Mail Bag No. 4 MOOROOKA QLD 4105
Phone:	(07) 3362 9494
Fax:	(07) 3362 9429
Email:	jessica.morgan@deedi.qld.gov.au
Objectives	To sequence two 800 base pair regions, one in the apicoplast genome (RpoC2 gene) and one in the mitochondrial genome (cytochrome b), for two strains of each of the seven species of chicken *Eimeria*.
Background	Coccidiosis, caused by any one of seven species of *Eimeria*, is costly for poultry industries around the world. Mixed species infections are common but overlapping sizes of oocysts make species difficult to distinguish morphologically. Current species identification uses molecular markers typically using assays based on nuclear DNA. Non-nuclear markers, such as mitochondrial and apicoplast DNA, offer independent genomes that carry a signature of vertically inherited divergence. With no recombination these genomes can acquire mutations more rapidly making them suitable for population genetic studies. To assess the utility of these two genomes at distinguishing among strains and species of *Eimeria*, a region of the RpoC2 gene was targeted in the apicoplast and the cytochrome b region of the mitochondrial DNA was sequenced and compared for two strains of each species of *Eimeria*.
Research	Primers were designed against publically available sequences from the *Eimeria tenella* genome project. DNA was extracted from pure lines of *Eimeria* species maintained at the Animal Research Institute in Brisbane. Target DNA was amplified using polymerase chain reaction and direct sequenced. Sequences were edited and aligned and phylogenetic trees were constructed to investigate relationships among the species and strains.
Outcomes	Sufficient variability was found in both the mitochondrial and apicoplast genetic markers to distinguish the seven described *Eimeria* species of chickens. No within-strain variability was detected making these genes well suited to diagnostic assays. Low level between-strain divergence was detected in four species suggesting that the non-nuclear genomes may contain enough variability to investigate strain differences. Phylogenetic analyses separate the strains into well supported species groups and indicate *E. tenella* and *E. necatrix* are sister species that branched early in the lineage.
Implications	The mitochondrial and apicoplast genomes will be suitable for developing diagnostic assays for the seven species of *Eimeria*. They may also contain enough mutations to distinguish among strains within a species.
Publications	Morgan JAT In prep The utility of non-nuclear genetic markers for species level diagnostics of *Eimeria* from chickens.