This is the peer-reviewed, manuscript version of an article published in *Veterinary Parasitology*. The version of record is available from the journal site: [https://doi.org/10.1016/j.vetpar.2018.03.020](https://doi.org/10.1016/j.vetpar.2018.03.020).

© 2018. This manuscript version is made available under the CC-BY-NC-ND 4.0 license [http://creativecommons.org/licenses/by-nc-nd/4.0/](http://creativecommons.org/licenses/by-nc-nd/4.0/).

The full details of the published version of the article are as follows:

**TITLE:** Discrimination, molecular characterisation and phylogenetic comparison of porcine Eimeria spp. in India

**AUTHORS:** Sharma, D., Singh, N. K., Singh, H., Joachim, A., Rath, S. S. and Blake, D. P.

**JOURNAL:** Veterinary Parasitology

**PUBLISHER:** Elsevier

**PUBLICATION DATE:** 22 March 2018 (online)

**DOI:** 10.1016/j.vetpar.2018.03.020
Discrimination, molecular characterisation and phylogenetic comparison of *Eimeria* spp. which infect pigs

Devina Sharma\textsuperscript{a,b,*}, Nirbhay K. Singh\textsuperscript{a}, Harkirat Singh\textsuperscript{a}, Anja Joachim\textsuperscript{c}, Shitanshu S. Rath\textsuperscript{a} and Damer P. Blake\textsuperscript{b}

\textsuperscript{a}Department of Veterinary Parasitology, College of Veterinary Science, Guru Angad Dev Veterinary and Animal Sciences University, Ludhiana – 141004, India

\textsuperscript{b}Department of Pathobiology and Population Sciences, Royal Veterinary College, North Mymms, Hertfordshire, University of London, United Kingdom AL97TA.

\textsuperscript{c}Institute of Parasitology, Department of Pathobiology, University of Veterinary Medicine, Veterinärplatz 1, A-1210 Vienna, Austria.

*Corresponding author. Email address: devinasharma23@yahoo.co.in (Devina Sharma)

Tel: +91 9418452080; Fax: +91 16102400822

Keywords

*Eimeria*, pig, 18S rDNA, India, *Sus scrofa domesticus*
ABSTRACT

Infections with *Eimeria* spp. are common in pigs worldwide, occasionally affecting animals clinically after weaning or during the fattening period when diarrhoea and weight loss can be observed upon infection with the more pathogenic species. Molecular characterization of pathogens is valuable to accurately delimit species and development novel diagnostics, although sequences which define *Eimeria* species that infect pigs are scarce. Only three of the eight common species are currently represented in GenBank. In this study we describe the occurrence of *Eimeria* species in pigs sampled in Punjab, India; going on to use the samples to generate new species-specific 18S rDNA sequences for all of the previously uncharacterised species. Using these data we report the first phylogenetic analyses to include the eight *Eimeria* species that commonly infect the domestic pig (*Sus scrofa domesticus*). Consideration of phylogenetic trees produced using Maximum Likelihood, Neighbour Joining, Maximum Parsimony and Unweighted Pair Group Method with Arithmetic Mean methods indicate that the 18S rDNA sequences present lower levels of genetic diversity than *Eimeria* which infect avian species and are insufficient to infer stable phylogenies.
1. INTRODUCTION

The genus *Eimeria* (Coccidia: Eimeriidae) includes numerous species, many of which can cause coccidiosis in birds and animals. In the domestic pig (*Sus scrofa domesticus*), 13 species of *Eimeria* have been described, eight of which are common (Pellerdy, 1974; Joachim and Schwarz, 2015). Infections occur worldwide and are usually caused by several species simultaneously (Vetterling, 1965; Löwenstein and Kutzer, 1989; Daugschies et al., 2004; Zhang et al., 2012; Gyzy and Oglu, 2016). Although most infections are asymptomatic, cases of diarrhoea, weight loss and even fatalities have been described in weaned pigs (Pellerdy, 1974). The species that have reportedly been involved in clinical eimeriosis are *Eimeria debliecki* (Vitovec and Koudela, 1990), *Eimeria scabra* (Rommel, 1970; Hill et al., 1985), *E. polita* (Rommel, 1970) and *Eimeria spinosa* (Koudela and Vitovec, 1992). Very high doses of *E. scabra* induced diarrhoea in weaned piglets after experimental infections (Pellerdy, 1974).

As pathogenic potential varies between *Eimeria* species, differentiation to the species level is important to determine those which contribute to disease in an outbreak or circulate on a farm. The eight most common porcine-infecting *Eimeria* spp, *E. debliecki*, *E. neodebliecki*, *E. perminuta*, *E. polita*, *E. porci*, *E. scabra*, *E. spinosa*, and *E. suis*, can readily be differentiated by oocyst morphology (Daugschies et al., 1999; Joachim and Schwarz, 2015), and algorithms can be applied to use morphometric analysis (Daugschies et al., 1999, 2004; Plitt et al., 1999). However, this requires careful preparation of the samples and suitable equipment which must be calibrated properly to avoid systemic errors (Oberg et al., 2013). Oocysts must be isolated from faeces and subjected to sporulation before measurement, which makes the processing cumbersome and unsuitable for larger sample numbers. There is thus a need for more rapid, accurate and cost effective identification tools, most notably in livestock other than poultry where the genus has been relatively neglected. Molecular characterisation can be applied to identify markers which are suitable for use in diagnostics and can improve understanding of evolutionary history and taxonomy of protozoan taxa (Ogedengbe et al., 2011). Molecular studies exploring sequence divergence and genetic distance have been useful to compare parasites isolated from different host species or geographic locations (Nadler and De Leon, 2011). Such genetic characterisation has benefitted our understanding of epidemiology, host and geographical ranges and evolutionary relationships. However, few sequences are publically available for *Eimeria* species which infect pigs. Several studies have compared 18S rDNA sequences from *Eimeria* species which infect chickens (Barta et al., 1997), turkeys (Miska et al., 2010), rodents (Zhao and Duszynski, 2001), rabbits (Kvicerová et al., 2008), goats (Nahavandi et al., 2016) and cattle (Kokusawa et al., 2013). Other genetic markers such
as the internal transcribed spacer (ITS) 1 and 2 sequences, and mitochondrial cytochrome
oxidase subunit 1 (mtCOI) have been sequenced from diverse Eimeria species (Ogedengbe et
al., 2011; Clark et al., 2016), although fewer sequences are publicly available at present. For
pigs 18S rDNA sequences were previously only available for three species, E. polita, E. porci
and E. scabra (Ruttkowski et al., 2001). The development of molecular diagnostics for Eimeria
which infect pigs requires a more comprehensive dataset. Here, we report partial 18S rDNA
sequences that represent each of the remaining five Eimeria species which commonly infect
pigs inferred from a series of mixed species samples collected in Punjab, India, by comparison
with oocyst morphology. We also validate those 18S rDNA sequences derived previously from
three Eimeria species which infect pigs.

2. MATERIALS AND METHODS

2.1 Sample collection

During August 2015 to September, 2016, a total of 839 faecal samples were collected
from 36 commercial and backyard pig farms covering 18 districts and 5 agro-climatic zones in
Punjab, North India (Fig. 1). The samples were collected from either the rectum at the time of
defaecation or the ground immediately after defaecation. All samples were transported to the
Department of Veterinary Parasitology, GADVASU, Ludhiana. Qualitative microscopic
analysis was done for Eimeria spp. by standard concentration-flotation technique and the
McMaster technique was used for the estimation of mean oocyst per gram (OPG) as per
Soulsby (1982). Those samples found to contain coccidial oocysts were fixed in 99.9% (v/v)
ethanol (Jebsen & Jessen, Germany) and transported to the Royal Veterinary College, UK
under Importation of Animal Pathogens Order (IAPO) permit for morphometric studies and
molecular analysis. Oocysts were separated from faecal debris and processed as described by
Kumar et al. (2014).

2.2 Ethical review

This study was reviewed and approved by the Dean of Post Graduate studies, Guru
Angad Dev Veterinary and Animal Science University (GADVASU), Ludhiana, Punjab, India.
More detailed ethical review was not required since no invasive sampling was undertaken.

2.3 Morphometric studies

Oocyst morphometry was assessed from each sample at 100x/400x using a dry high
power objective with a photomicrographic camera (Olympus CX41) attached to a binocular
research microscope (Olympus). Micrometric analysis was performed using Image J (Austen et al., 2014) and species identity was assigned using morphological features of the oocysts including shape, size, colour, shape index and micropyle as described elsewhere (Daugschies et al., 1999; Ramirez et al., 2008).

2.4 DNA extraction

Total genomic DNA was extracted as described previously (Kumar et al., 2014). Briefly, Aliquots of ~0.75ml from samples found to contain more than 500 oocysts/ml were sedimented by centrifugation at 10,000 × g for 1min and resuspended in 0.5 ml phosphate buffered saline (PBS, pH 7.6). Sedimentation was repeated prior to mechanical disruption using 0.4-0.6mm glass beads (Sigma-Aldrich, St Louis, USA) and a Bead Beater at 30,000 × oscillations/min for 30s. Total genomic DNA was then extracted using a QIAamp fast DNA Stool Mini kit as described by the manufacturer (Qiagen, Hilden, Germany). Eluted DNA quality and quantity was determined using a Nanodrop 2000/200C spectrophotometer (Thermo Fisher Scientific, USA) prior to storage at −20 °C until use.

2.5 Polymerase chain reaction, gel electrophoresis and amplicon sequencing

Initially, ~1,800 bp fragments of the 18S rDNA were amplified in 50 μl PCR reactions including 4μl (2-20 ng/μl) genomic DNA template, 25 pM of the primers ER1B1 (5’-ACCTGGTGTACCCGCAAGTG-3’) and ER1B10 (5’-CTTTCGAGTTACCTACTGGGGG-3’) (Schwarz et al., 2009) and 25 μl of 2× MyTaq™ Mix (Bioline, Taunton, USA), made up with DNase/RNase free molecular grade water (ThermoFisher Scientific, Hemel Hempstead, UK). PCR amplification was initiated at 95ºC for 2 min followed by 40 cycles of 95ºC for 30sec, 56ºC for 30 sec (adjusted to optimise efficiency from 57ºC in Schwarz et al., 2009), 72ºC for 2 min and a final elongation step at 72ºC for 10 min. Subsequently, primers Pig18S_F1 (5’-GGATTTCGTGTCGCATGCTAC-3’) and Pig18S_R1 (5’-CTTTAAGTTCAGCTTGGG-3’) were designed to amplify ~510-535 bp fragments of the porcine-infecting Eimeria 18S rDNA flanking the region of greatest variation for medium depth amplicon sequencing. *Eimeria tenella* genomic DNA and molecular grade water served as template in positive and no template negative controls.

PCR amplicons were resolved through 1.0% (w/v) Ultrapure Agarose (Invitrogen, Paisley, UK) in 0.5× TBE (0.89 M Tris base, 0.89 M boric acid, 0.5 M ethylenediaminetetraacetic acid [EDTA] buffer; Sigma-Aldrich), including 0.01% (v/v) SafeView Nucleic Acid Stain (Novel Biological Solutions, Huntingdon, UK). 5 μl of each
amplicon was mixed with 1 μl of 6× DNA Loading Dye (ThermoFisher Scientific) and then subjected to electrophoresis at 60V/30 min. Amplicons of the anticipated size were purified using a QIAquick® PCR Purification Kit (Qiagen), as per the manufacturer’s instructions.

Purified PCR amplicons were cloned using pGEM-T Easy (Promega) in XL1-Blue MRF Escherichia coli (Stratagene) as described by the manufacturers, and screened for inserts of the anticipated size by colony PCR using the same primers and conditions as described for the original amplification. Those plasmids with confirmed inserts were purified using a QIAprep® Spin Mini prep kit (Qiagen, Germany) and sequenced using the T7 and SP6 T Easy sequencing primers (GATC Biotech, Konstanz, Germany) as described by the respective manufacturers. Each field sample screened included multiple Eimeria species (based upon morphometry). Thus, up to ten clones were sequenced for each sample, assigning sequence identity by comparison of the proportionate occurrence of (i) the dominant Eimeria species morphotype and (ii) the most frequent sequence type, as described previously for Eimeria which infect cattle (Kawahara et al., 2010).

2.6 Sequence analysis and phylogenetics

Newly generated and existing 18S rDNA sequences downloaded from GenBank were aligned using CLC Main Workbench v.6.9.1 (CLC bio, Aarhus, Denmark). The Basic Local Alignment Search Tool (BLAST®; http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to determine similarity with existing sequence resources. The percentage identity/mean genetic distance between sequences was calculated with software MEGA 6 (Tamura et al., 2013), including reference sequences for E. polita, E. porci and E. scabra (GenBank accession numbers AF279666-8) as well as sequences representing members of the Eimeriidae isolated from mice, chickens, cattle, sheep and turkeys, employing Toxoplasma gondii as a taxonomic and functional outgroup (Supplementary Table 1). Phylogenies were inferred using Maximum likelihood (ML), Neighbor Joining (NJ), Maximum Parsimony (MP) and Unweighted Pair Group Method with Arithmetic Mean (UPGMA) methods with MEGA 6 (Tamura et al., 2013), including 1,000 bootstrap replication. The evolutionary distances were computed using the Kimura 2 parameter. The rate variation among sites was modelled with a gamma distribution (shape parameter = 1). The branching order was confirmed by nucleotide substitution model of p distance with the resampling nodal support of 1,000 bootstrap replicates, under the Maximum Parsimony (MP) analysis using the close-neighbour-interchange algorithm. Mean genetic distance was calculated using MEGA 6 with default parameters.
3. RESULTS

3.1 Description and frequency of *Eimeria* spp.

A panel of 839 faecal samples from 36 commercial and backyard pig farms were collected from Punjab, North India. Farms were selected for sampling using the Progressive Pig Farmers Association of Punjab database following a semi-randomised process, modified for convenience of access. Faecal samples were collected from individual animals. Microscopic examination identified 79/839 (9.4%) samples were positive for coccidia. Out of 36 farms screened, 23 (63.9%) showed the presence of *Eimeria* oocysts with a minimum of two species, a maximum of seven, and most farms having six species (Supplementary Fig. 1). Total oocyst counts ranged from 1,000-62,000 per ml (Supplementary Table 2). From the 79 positive samples 55.7% were collected from piglets, 20.3% from growers 24.1% from adults, indicating that older pigs also contributed to oocyst shedding. The differences between age groups were not statistically significant (Pearson chi square test, p=0.8). Morphometric analysis identified examples of all eight *Eimeria* species (*E. debliecki*, *E. neodebliecki*, *E. perminuta*, *E. polita*, *E. porci*, *E. scabra*, *E. spinosa* and *E. suis*).

3.2 Polymerase chain reaction and molecular cloning

Initially, 19 samples were selected for DNA extraction based on either high oocyst occurrence or low complexity of infection. Primers ERIB1/ERIB10 were optimised to confirm the successful extraction of sufficient eimerian genomic DNA from each sample, indicating 19/19 samples were appropriate for further PCR analysis. Subsequently, custom primers flanking the most variable region of the porcine-infecting *Eimeria* 18S rDNA were designed to amplify ~510-535 bp to optimize cloning efficiency and plasmid-based Sanger sequencing. Samples found by morphometry to contain (i) only two *Eimeria* species or (ii) one dominant species were selected for PCR and cloning to allow unambiguous allocation of genotype to oocyst morphotype. In total 31 new sequences were generated from 13 samples, all of which are available from GenBank under the accession numbers LT962626-56. Subtracting the published 18S rDNA sequences for *E. polita*, *E. porci* and *E. scabra* (AF279666-8), we were able to identify five novel sequences which were related to *E. debliecki*, *E. neodebliecki*, *E. perminuta*, *E. spinosa* and *E. suis* (LT962626-30), as well as a second sequence type which was associated with *E. polita* (LT962633, termed *E. polita* sequence type (ST) II). These putative new reference sequences were then used to annotate the remaining sequences.

3.3 Phylogenetic analysis
ML, NJ, MP and UPGMA methods were applied to the reference and new 18S rDNA sequences, with the ML and NJ versions shown here (Fig. 2A & B). Comparable results were obtained using all four methods with consistent topology between Eimeria species, although bootstrap values at nodes between Eimeria sequences derived from pigs indicated limited support for many. Comparison of BLAST sequence identity and all phylogenetic trees indicated nine clusters, each defined by a single published or new reference sequence. Cluster I contained four sequences and were characterised as E. debliecki. Cluster II contained four sequences characterised as E. perminuta. Clusters III, V, VI, VIII and IX represented E. neodebliecki (n=5), E. porci (n=3), E. suis, E. spinosa and E. scabra (all n=2), respectively. Clusters IV and VII represented the two E. polita sequence types, with four and 11 sequences respectively (ST II and I). A close relationship was indicated for E. porci and E. suis. Mean genetic diversity within each sequence cluster ranged from 0.000 to 0.015 (Table 2). When analysed in the context of other publically available Eimeria sequences, the pig-Eimeria specific partial 18S rDNA sequences formed a monophyletic cluster distant to those species which infect chickens, turkeys and ruminants such as cattle and sheep. Mean genetic diversity within the sequences generated from Eimeria which infect pigs was lower than calculated for Eimeria which infect chickens, turkeys or sheep (Table 2).

4. DISCUSSION

Thirteen Eimeria species have been described and at least nine are valid. Here, all eight of the most common Eimeria species known to infect domestic pigs were detected during analysis of faecal samples collected from pigs in Punjab, North India. These Eimeria species are ubiquitous and have been widely reported (Vetterling, 1965; Pellérdy, 1974; Löwenstein and Kutzer, 1989; Chhabra and Mafukidze, 1992; Roepstorff et al., 1998; Lai et al., 2011; Tsunda et al., 2013; Joachim and Schwarz, 2015). Pig age did not significantly influence the occurrence of Eimeria infection. The oocyst lifecycle stage of each species is phenotypically distinct, although overlap in features such as length and width, make routine species-level identification beyond the differentiation challenging (Daugschies et al., 1999). Prior to this work just three sequences derived from porcine-infecting Eimeria were publicly available, published in a single study (Ruttkowski et al., 2001). The scarcity of sequence resources for Eimeria which infect pigs has limited the production of novel molecular diagnostics and precluded epidemiological investigation.

Here, we have targeted the 18S rDNA of the genus Eimeria (El-Sherry et al., 2013). The 18S rDNA is the only genomic locus to have been sequenced previously for Eimeria which
infect pigs. In the absence of single-species DNA samples such prior data was invaluable to untangle the complexity presented by the field samples available to the project. Other benefits included the presence of multiple 18S copies in each parasite genome, improving sensitivity, and the presence of highly conserved flanking regions appropriate for universal primers (Kokusawa et al., 2013). Here, we sequenced multiple clones derived from PCR amplicons representative of the least complex field samples. Through comparison with species occurrence in each sample, and by subtraction of known sequences as they became available, a series of nine sequence clusters were identified. Comparison with morphometric data permitted identification of sequences specific for each *Eimeria* species, with two sequence types associated with *E. polita*. Multiple 18S rDNA sequence types have previously been described for other apicomplexans such as *E. mitis* from the chicken (Vrba et al., 2011). Alternatively, the one of the two sequence types might represent another *Eimeria* species. The nine sequence clusters associated with porcine *Eimeria* formed a monophyletic lineage divergent from those which infect chickens, turkeys and ruminants, proving closer to *E. falciformis* which infects mice. The use of a single genetic locus precludes a firm conclusion, but the data does indicate distinct parasite lineages associated with avians, ruminants and monogastrics. Sequence assignment to each species was largely consistent between phylogenetic methods, although bootstrap confidence was consistently low. Thus, delimitation of species on the basis of sequence congruence and phylogeny was difficult in the absence of defined reference specimens appropriate for more extensive sequencing (Ruedas et al., 2000). Miska and colleagues have previously reported similarly challenging analysis using 18S and mtCOI sequences where the allocation of species identity was limited in mixed *Eimeria* samples from turkeys, pheasants and chukar (Miska et al., 2010). Nuclear 18S rDNA sequences have occasionally been unable to resolve the monophyly of *Eimeria* (Morrison et al., 2004) and is perhaps only reliably useful in phylogenetic analyses to the genus level (El-Sherry et al., 2013).

Mean genetic distances within and between *Eimeria* species which infect pigs were lower than have been described elsewhere for other *Eimeria* species (Blake et al., 2015; Clark et al., 2016; Clark et al., 2017). This might have been an artefact of the narrow spatial sampling range, although higher levels of diversity have been reported in *Eimeria* which infect chickens in a similar region (Blake et al., 2015). Other possibilities might include limited opportunities for transmission of novel genotypes or cross-fertilization as a consequence of the relatively small pig population in Punjab, although a broader panel of genetic markers will be required to answer such questions.
A major objective of these studies was to develop a panel of PCR-based *Eimeria* species-specific diagnostic tests. However, the level of inter-specific diversity discovered was insufficient. Moreover, comparison of the expanded sequence resources demonstrated that the putative species-specific assays developed previously using monospecific *E. polita, E. porci* and *E. scabra* samples (Ruttkowski et al., 2001) are not, in fact, specific. Application of these primers here revealed a considerable number of false positive results (data not shown). Thus, while direct species-specific diagnostics are not yet feasible robust genus-specific tools are realistic with the addition of sequencing required for species identification (Nahavandi et al., 2016). Access to pure isolates of each *Eimeria* species would permit the application of next-generation sequencing technologies to expand molecular resources for these parasites as suggested for coccidia of other host species (Lim et al., 2012; Diaz-Sanchez et al., 2013; Vermeulen et al., 2016). As single cell genomics technologies become more effective such resources may soon become available, with direct relevance to the development of novel diagnostics and innovative control strategies.

**ACKNOWLEDGEMENTS**

The authors would like acknowledge Guru Angad Dev Veterinary and Animal Science University and the Commonwealth Scholarship Commission for funding through a Split Site PhD Scholarship. The RVC have assigned this manuscript the reference PPS_01741.

**REFERENCES**


Hill, J.E., Lomax, L.G., Lindsay, D.S., Lynn, B.S. 1985. Coccidiosis caused by Eimeria


Nahavandi, K.H., Mahvi, A.H., Mohebali, M., Keshavarz, H., Rezaei, S., Mirjalali, H,


Soulsby E J L. 1982. Helminths, Arthropod and Protozoa of Domesticated Animals, 7th edn,


Table 1. The number and age of pigs sampled in this study, indicating the number found to have been excreting detectable numbers of coccidial oocysts. *No significant difference was found between age groups (Pearson chi square test, p=0.8).

<table>
<thead>
<tr>
<th>Age</th>
<th>Examined (n)</th>
<th>Positive (%)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piglet(&lt;4m)</td>
<td>341</td>
<td>44 (12.9)</td>
</tr>
<tr>
<td>Grower(4-8m)</td>
<td>228</td>
<td>16 (7.0)</td>
</tr>
<tr>
<td>Adult(&gt;8m)</td>
<td>270</td>
<td>19 (7.0)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>839</strong></td>
<td><strong>79 (9.4)</strong></td>
</tr>
</tbody>
</table>

Table 2. Mean genetic distance within (i) each individual *Eimeria* species which infects pigs and (ii) all *Eimeria* species which infect pigs, chickens, sheep or turkeys. ST = sequence type.

<table>
<thead>
<tr>
<th>Genus/species</th>
<th>Mean genetic distance</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. debliecki</em></td>
<td>0.0150</td>
</tr>
<tr>
<td><em>E. neodebliecki</em></td>
<td>0.0064</td>
</tr>
<tr>
<td><em>E. perminuta</em></td>
<td>0.0081</td>
</tr>
<tr>
<td><em>E. polita</em> (ST I)</td>
<td>0.0044</td>
</tr>
<tr>
<td><em>E. polita</em> (ST II)</td>
<td>0.0040</td>
</tr>
<tr>
<td><em>E. porci</em></td>
<td>0.0095</td>
</tr>
<tr>
<td><em>E. scabra</em></td>
<td>0.0144</td>
</tr>
<tr>
<td><em>E. spinosa</em></td>
<td>0.0000</td>
</tr>
<tr>
<td><em>E. suis</em></td>
<td>0.0000</td>
</tr>
</tbody>
</table>

*Eimeria* (All, pig) 0.0130
*Eimeria* (All, chicken) 0.0375
*Eimeria* (All, sheep) 0.0276
*Eimeria* (All, turkey) 0.0291
Figure 1. A Maximum Likelihood. B Neighbor Joining.