

# Population, genetic, and antigenic diversity of the apicomplexan *Eimeria tenella* and their relevance to vaccine development

Damer P. Blake<sup>a,1</sup>, Emily L. Clark<sup>a,2</sup>, Sarah E. Macdonald<sup>a</sup>, Venkatachalam Thenmozhi<sup>b</sup>, Krishnendu Kundu<sup>c</sup>, Rajat Garg<sup>c</sup>, Isa D. Jatou<sup>d</sup>, Simeon Ayoade<sup>e</sup>, Fumiya Kawahara<sup>f</sup>, Abdalgader Moftah<sup>g</sup>, Adam James Reid<sup>h</sup>, Ayotunde O. Adebambo<sup>e</sup>, Ramón Álvarez Zapata<sup>i</sup>, Arni S. R. Srinivasa Rao<sup>j,k</sup>, Kumarasamy Thangaraj<sup>l</sup>, Partha S. Banerjee<sup>c</sup>, G. Dhinakar-Raj<sup>m</sup>, M. Raman<sup>b</sup>, and Fiona M. Tomley<sup>a</sup>

<sup>a</sup>Pathology and Pathogen Biology, Royal Veterinary College, North Mymms, Hertfordshire, AL9 7TA, United Kingdom; <sup>b</sup>Department of Veterinary Parasitology, Madras Veterinary College, Tamil Nadu Veterinary and Animal Sciences University, 600 007, Chennai, India; <sup>c</sup>Division of Parasitology, Indian Veterinary Research Institute, Izatnagar, 243 122, Uttar Pradesh, India; <sup>d</sup>Department of Parasitology and Entomology, Faculty of Veterinary Medicine, Ahmadu Bello University, Zaria, Nigeria; <sup>e</sup>Department of Animal Breeding and Genetics, Federal University of Agriculture, Abeokuta, Ogun State, Nigeria; <sup>f</sup>Nippon Institute for Biological Science, Tokyo 198-0024, Japan; <sup>g</sup>School of Agriculture, Food and Rural Development, Newcastle University, Newcastle upon Tyne, NE1 7RU, United Kingdom; <sup>h</sup>Parasite Genomics, Wellcome Trust Sanger Institute, Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, United Kingdom; <sup>i</sup>Facultad de Agronomía Instituto de Producción Animal, Central University of Venezuela, Maracay, Venezuela; <sup>j</sup>Department of Biostatistics and Epidemiology Georgia Regents University, Augusta, GA 30912; <sup>k</sup>Department of Mathematics, Georgia Regents University, Augusta, GA 30912; <sup>l</sup>Council of Scientific and Industrial Research-Centre for Cellular and Molecular Biology, Hyderabad 500 007, India; and <sup>m</sup>Department of Animal Biotechnology, Madras Veterinary College, Tamil Nadu Veterinary and Animal Sciences University, Chennai 600 007, India

Edited by Jitender P. Dubey, US Department of Agriculture, Beltsville, MD, and approved August 21, 2015 (received for review April 1, 2015)

The phylum Apicomplexa includes serious pathogens of humans and animals. Understanding the distribution and population structure of these protozoan parasites is of fundamental importance to explain disease epidemiology and develop sustainable controls. Predicting the likely efficacy and longevity of subunit vaccines in field populations relies on knowledge of relevant preexisting antigenic diversity, population structure, the likelihood of coinfection by genetically distinct strains, and the efficiency of cross-fertilization. All four of these factors have been investigated for *Plasmodium* species parasites, revealing both clonal and panmictic population structures with exceptional polymorphism associated with immunoprotective antigens such as apical membrane antigen 1 (AMA1). For the coccidian *Toxoplasma gondii* only genomic diversity and population structure have been defined in depth so far; for the closely related *Eimeria tenella*, a major cause of the enteric disease coccidiosis, which exerts a profound effect on chicken productivity and welfare, we determined population structure, genotype distribution, and likelihood of cross-fertilization during coinfection and also investigated the extent of naturally occurring antigenic diversity for the *E. tenella* AMA1 homolog. Using genome-wide Sequenom SNP-based haplotyping, targeted sequencing, and single-cell genotyping, we show that in this coccidian the functionality of *EtAMA1* appears to outweigh immune evasion. This result is in direct contrast to the situation in *Plasmodium* and most likely is underpinned by the biology of the direct and acute coccidian life cycle in the definitive host.

*Eimeria* | coccidiosis | population structure | chickens | food security

The phylum Apicomplexa contains protozoan parasites of great medical and veterinary importance including *Plasmodium*, *Toxoplasma*, and *Eimeria* species. Cost-effective vaccines and/or drugs are urgently required to control the diseases caused by these pathogens, but their complex life cycles and naturally occurring genetic polymorphism makes the development of such vaccines an extremely difficult task. For *Plasmodium* species and *Toxoplasma gondii*, in-depth studies have defined genetic diversity and population structures in many regions and populations, with descriptions ranging from clonality to panmixia (e.g., refs. 1–3). For *Eimeria* it is widely perceived that species occurrence has been established and that the risks posed are well understood, but population structures within field populations are unclear, and the occurrence of genetic diversity is unexplored (4). Given the recent progress in identifying antigens for inclusion in novel subunit vac-

cines to control poultry coccidiosis (reviewed in ref. 5), it is now timely to investigate the extent, distribution, and dissemination of genetic diversity in this parasite.

Sustainable food security is a major concern with the global human population set to exceed nine billion by 2050 (6). Poultry are the most efficient source of animal-derived protein (7), and consequentially pathogens that compromise poultry production are reemerging as threats to the global food supply and as factors in human poverty (8, 9). Coccidiosis is a problem in many livestock systems; most notably, in chicken (broiler) production seven different species of *Eimeria* cause serious intestinal disease

## Significance

Sixty billion chickens are produced worldwide each year, and all are at risk from *Eimeria*, parasites that cause coccidiosis. Control relies widely on chemoprophylaxis, but pressure to reduce drug use in farming urges development of cost-effective vaccines. Antigens such as apical membrane antigen 1 (AMA1) offer promise as anticoccidial vaccine candidates, but experience with related apicomplexans such as *Plasmodium*, in which pre-existing antigenic diversity and incompatible population structure have undermined vaccine development, tempers confidence. Parasite genotyping identified enormous region-specific variation in haplotype diversity for *Eimeria tenella* but a contrastingly low level of polymorphism for *EtAMA1*. Although high levels of polyclonal *Eimeria* infection and hybridization indicate an ability to disseminate vaccine resistance rapidly, the low level of *EtAMA1* diversity promotes vaccine development.

Author contributions: D.P.B., E.L.C., K.T., P.S.B., G.D.-R., M.R., and F.M.T. designed research; D.P.B., E.L.C., S.E.M., V.T., K.K., R.G., I.D.J., S.A., F.K., A.M., A.J.R., A.O.A., R.Á.Z., and A.S.R.S.R. performed research; V.T., K.K., R.G., I.D.J., S.A., F.K., A.M., A.O.A., R.Á.Z., P.S.B., G.D.-R., and M.R. contributed new reagents/analytic tools; D.P.B. and E.L.C. analyzed data; and D.P.B., A.S.R.S.R., K.T., and F.M.T. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. LN609976–LN610031 and LN626984–LN626991).

<sup>1</sup>To whom correspondence should be addressed. Email: dblade@rvc.ac.uk.

<sup>2</sup>Present address: The Roslin Institute, The University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, United Kingdom.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1506468112/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1506468112/-DCSupplemental).

that compromises economic productivity and animal welfare, incurring global costs in excess of \$3 billion every year (10, 11). Disease control relies on the use of in-feed drugs, but multidrug resistant parasites are ubiquitous, and concerns about residues are widespread (12). Live anticoccidial vaccines have been used in small sectors of the poultry industry for more than 50 y with no evidence of parasite evolution toward resistance/immune escape (11). Examples of naturally occurring genetic diversity have been reliably overcome by the inclusion of more than one “strain” of a single species in some vaccine formulations (11). Each *Eimeria* species expresses between 6,000 and 9,000 proteins throughout its lifecycle (13), exposing the host to a complex portfolio of antigens. Genetic mapping studies with *Eimeria maxima* indicated that at least six loci associate absolutely with susceptibility to strong natural immune responses, and additional loci under partial selection also were detectable (14). This breadth of antigenic repertoire is likely to have limited the impact of diversifying selection on the efficacy of live vaccines and explains why resistance has not developed. Unfortunately, these live vaccines must be propagated in chickens, and the low reproductive index of attenuated vaccine parasites makes production costs high and limits capacity (11). More cost-effective subunit vaccines composed of a small number of *Eimeria* antigens are needed for the mass broiler market; however, these vaccines are likely to induce more focused immune selection, which potentially could drive rapid selection of vaccine-resistant populations. *Eimeria tenella* is considered one of the most important species because it is widespread and highly pathogenic (13, 15). It also is the best-studied species and therefore was selected as the exemplar in these studies. We used PCR to define the occurrence of *E. tenella* within fecal samples collected from domestic chickens (*Gallus gallus domesticus*) across five continents. Positive samples then were genotyped using a custom Sequenom MassARRAY SNP tool, allowing us to define population structure for an *Eimeria* species for the first time, to our knowledge.

## Results

### Description of the Farms Sampled and the Occurrence of *E. tenella*.

Fecal samples were collected from 595 small- to medium-scale commercial chicken farms (defined here as holding up to 50,000 birds) from eight countries, supplemented by laboratory reference strains and, when combined, representing five continents. Broiler and layer birds were sampled in approximately equal numbers in each region. Each sample comprised several fecal pellets collected from a cage or pen; thus parasites were derived from multiple birds within a single “flock.” PCR screening found *E. tenella* to be present in all regions that were sampled more than once, most frequently in Nigeria and Southern India (Andhra Pradesh, Karnataka, Kerala, and Tamil Nadu combined), (>80% of farms) and less frequently in Northern India (Haryana, Punjab, Uttarakhand, and Uttar Pradesh combined), Egypt, and Libya (37–66%) (Fig. S1 and Table S1).

### *E. tenella* Is Genetically Diverse with Evidence of Allopatric Evolution.

Sequence-led studies of genetic diversity within *Eimeria* species most commonly have focused on the internal transcribed spacer (ITS) sequences 1 and 2 (4). The large number of ITS copies per genome provides a high degree of sensitivity by PCR, although the occurrence of polymorphism between copies, even within single genomes, can result in overestimation of clonal diversity. For more detailed comparisons, we investigated genome-wide genetic diversity among *E. tenella* samples. First, genetic polymorphism between the Houghton (United Kingdom) reference and a panel of archive *E. tenella* strains was assessed using amplified fragment-length polymorphism (AFLP) to calculate Jaccard indices of diversity in the absence of more extensive genomic resources (13, 16). Comparison of the Houghton strain

with the LCH2 and Weybridge (United Kingdom.), Beltsville and Wisconsin (United States), and Nippon-2 (Nt2; Japan) *E. tenella* strains revealed Jaccard indices of 0.912, 0.967, 0.956, 0.914, and 0.873, respectively. Using these data, we selected the Wisconsin and Nt2 strains for next-generation sequencing (Illumina), as described previously (13). The sequences are available from the European Nucleotide Archive under the accession number PRJEB4009. Using these resources, we identified SNPs. First, sequences covering a panel of genes with known chromosomal locations were selected to provide an even dispersal of SNPs across the 14 *E. tenella* chromosomes (Dataset S1) (13, 17). Additional SNPs were selected across chromosomes 1 and 2, including two to four SNPs in each repeat-rich (R) and repeat-poor (P) region (Dataset S1) (13, 18). In total, 55 SNPs were selected, incorporated into two Sequenom MassARRAY multiplexes, and used to genotype 244 field samples found to be PCR positive for *E. tenella* (Table 1). After SNP genotyping 52 of the 55 SNPs produced informative profiles.

For population genetic analysis, SNP profiles were presented as the major genotype in each sample, determined by the Sequenom SNP frequency of each allelic type in cases of polyclonal populations. Where polyclonal populations were detected the major and minor SNP types were identified, and the major allele was selected for use in haplotype analysis without phasing, because each parasite presented a haploid genome at the time of sampling. The lack of replication per sample and the semi-quantitative nature of the assay precluded computational assembly of haplotype complexity within each sample (19). Analysis of all SNP genotypes using DnaSP v5.10.01 (20) identified 93 SNP haplotypes in total, with striking variation in haplotype depth and diversity among regions (Table 1). In Southern India 50 SNP haplotypes were detected from 53 samples, whereas in Northern India only eight were identified from 86 samples ( $P < 0.005$ ;  $\chi^2$  test). Comparison between regions revealed that 98% and 87.5% of these haplotypes were unique to their regions, providing evidence of allopatric diversification. In contrast, Egypt and Libya presented an overlapping haplotype profile, so that pooling data for these neighboring countries revealed a similar level of regional specificity. Complementary analysis using STRUCTURE v2.3.4 (21) to assess population

**Table 1. Summary of *E. tenella* genome-wide genetic data calculated using Sequenom MassARRAY genotyping**

| Region         | N  | SNP haplotypes |                 | Het ( $\pm$ )   | $I_A^S$ | P       |
|----------------|----|----------------|-----------------|-----------------|---------|---------|
|                |    | Total          | Region specific |                 |         |         |
|                |    | All            | 93              |                 |         |         |
| Northern India | 86 | 8              | 7               | 0.3331 (0.0254) | 0.1153  | <0.0001 |
| Southern India | 53 | 50             | 49              | 0.3405 (0.0212) | 0.0087  | ns      |
| Egypt          | 40 | 21             | 13              | 0.1925 (0.0316) | 0.0597  | <0.0001 |
| Libya          | 51 | 11             | 3               | 0.1611 (0.0247) | 0.0533  | <0.0001 |
| Egypt + Libya  | 91 | 25             | 24              | 0.1789 (0.0275) | 0.0533  | <0.0001 |
| Nigeria        | 14 | 14             | 13              | 0.3998 (0.0224) | 0.0012  | ns      |

Northern India included samples collected in Haryana, Punjab, Uttarakhand, and Uttar Pradesh. Southern India included samples collected in Andhra Pradesh, Karnataka, Kerala, and Tamil Nadu. N, the number of samples tested. The number of SNP haplotypes was calculated using DnaSP v5.10.01. Het, mean genetic diversity;  $I_A^S$ , standardized index of association (calculated using only the first SNP per specific gene-associated sequence or chromosomal repeat type segment to avoid bias as a consequence of close physical linkage), with statistical significance indicated as P, both calculated using LIAN v3.6; ns, not statistically significant.

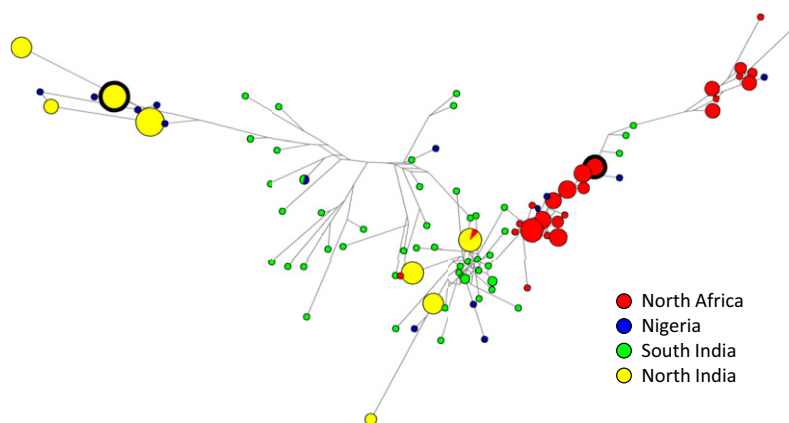
structure ( $K$ ) suggested that eight ancestral genotype clusters best supported the current data, with log likelihood [ $\ln P(D)$ ] plateauing when  $K = 7$ , variation increasing when  $K = 9$ , and the mean difference between successive likelihood values of  $K$  [ $L'(K)$ ] peaking at  $K = 9$  (Fig. S2 A and B). STRUCTURE output for  $K$  values of 7, 8, and 9 are shown in Fig. S2 C–E, and the input data are provided in Dataset S2. Representation of the data using NETWORK v4.6.1.1 (22) to calculate a median-joining method phylogenetic network illustrated these findings (Fig. 1). A small number of large nodes were calculated for Northern India and North Africa (data from Egypt and Libya combined) distributed across a small number of clusters, indicating multiple occurrences of conserved haplotype profiles. In contrast, Southern India presented a much larger number of smaller nodes in loosely defined clusters, representing the greater haplotype diversity and narrow population structure. The small number of Nigerian samples analyzed revealed the greatest genetic diversity with a broad spread across much of the phylogenetic network (Fig. 1 and Table 1). The distance between clusters was confirmed by calculating the mean genetic diversity (Het) within each region using LIAN v3.6 (23); the regions with more widely distributed clusters presented higher Het (Table 1).

Linkage disequilibrium (LD) was calculated within the full and regional subpopulations using LIAN v3.6 to test the null hypothesis of linkage equilibrium for all loci. In recognition of the possible bias introduced by SNPs with a close physical linkage, we used the standardized index of association ( $I_A^S$ ) using a restricted SNP panel with only the first SNP in each gene-associated sequence or chromosome-specific repeat-type region (the SNPs used are shown in Dataset S1). Statistically significant LD was identified in Northern India and North Africa but not Southern India or Nigeria (Table 1), suggesting that mating and recombination occur freely in the latter but may be restricted in the former populations.

**Genomic Diversity Is Associated with Limited Escape from Strain-Specific Immune Killing for *E. tenella*.** To investigate the likely efficacy of the current generation of anticoccidial vaccines against genetically diverse *E. tenella*, we tested cross-protective ability between two of the most distantly related strains available. The Houghton reference strain, isolated in the United Kingdom and progenitor for the attenuated *E. tenella* line included in the Paracox vaccine series, is most closely related to the Libyan LI\_01 NETWORK cluster (red node with heavy black boundary in Fig. 1). As a comparison, we chose a clonal line isolated from Northern Indian sample NI\_01 (24), which has one of the most

distant SNP NETWORK profiles (yellow node with heavy black boundary in Fig. 1). Initially, immunization of outbred Light Sussex chickens at 3 and 4 wk of age used sequential doses of 1,000 and 4,000 sporulated oocysts of either the Houghton or NI strains. Subsequent homologous challenge with 250 sporulated oocysts 2 wk later was blocked, with no detectable oocyst output (Fig. S3A). Heterologous challenge was similarly blocked completely when the Houghton strain was used to challenge birds immunized by the NI strain. However, when the NI strain was used to challenge birds immunized by the Houghton strain, there was a low level of escape, with birds excreting 0.6% of the oocysts produced by unimmunized NI challenge controls. Repeating the trial for only the NI challenge component with a lower-level immunization protocol (sequential doses of 100 and 1,000 sporulated oocysts per bird, administered at 3 and 4 wks of age, respectively) revealed incomplete homologous protection (12.7% oocysts produced compared with the unimmunized control) and a significantly greater escape following heterologous challenge (34.8%;  $P < 0.05$ , ANOVA plus Bonferroni post hoc test) (Fig. S3B).

**Sequence Diversity Is Limited for the Anticoccidial Vaccine Candidate Apical Membrane Antigen 1.** Experimental testing of subunit anticoccidial vaccines containing recombinant antigens such as apical membrane antigen-1 (AMA1) has prompted interest in the extent of naturally occurring genetic diversity in these antigens and has raised the question of whether polymorphism might result in vaccine breakthrough by field parasite populations (14). We sequenced and assembled full-length AMA1-coding sequences for 55 *E. tenella* samples recovered from four continents (accession nos. LN609976–LN610031) and compared these sequences with the published sequence JN032081 (Table 2 and Table S1). Analysis of the 1,608-bp coding sequence alignment revealed only 13 nucleotide substitutions and one short tandem repeat (STR) variation of the trinucleotide CTG, which is known to be common among *Eimeria*-coding sequences (13). Ten substitutions were nonsynonymous, resulting in nine amino acid substitutions (two were located within codon 292, glutamine to glycine; Fig. 24). The STR variation incurred expansion/reduction of a leucine homopolymeric amino acid repeat. The greatest diversity in *ama1* was in samples from Asia (China, India, and Japan) and Nigeria as defined by average pairwise difference, nucleotide diversity calculated as  $\pi$  with the Jukes Cantor correction, haplotype number, and haplotype diversity using DnaSP v5.10.01 (Table 2) (20). In contrast to the SNP-based



**Fig. 1.** Median-joining phylogenetic NETWORK showing the relationships between Sequenom MassARRAY haplotypes for *E. tenella* samples collected in Northern India (yellow), Southern India (green), Egypt and Libya combined (North Africa, red), and Nigeria (blue). Node size indicates the frequency of haplotype occurrence. Nodes surrounded by a heavy black boundary indicate the haplotypes used in the cross-protection studies.

**Table 2. Summary of AMA1 genetic data**

| Region               | N  | S       | dN | dS | k     | $\pi$   | Jukes Cantor | H     | Hd | H <sup>†</sup> | Tajima's D         | Fu and Li's D* | Fu and Li's F* |
|----------------------|----|---------|----|----|-------|---------|--------------|-------|----|----------------|--------------------|----------------|----------------|
| All                  | 56 | 13 (13) | 10 | 3  | 4.041 | 0.0032  | 8            | 0.771 | —  | 1.259          | 1.513**            | 1.688**        |                |
| Asia                 | 7  | 12 (12) | 9  | 3  | 6     | 0.00475 | 4            | 0.857 | 2  | 1.225          | 1.573 <sup>†</sup> | 1.642**        |                |
| Egypt/Libya          | 7  | 3 (3)   | 0  | 3  | 1.714 | 0.00135 | 2            | 0.571 | 0  | 1.811          | 1.297              | 1.522          |                |
| Europe/United States | 6  | 4 (4)   | 1  | 3  | 2.4   | 0.00189 | 3            | 0.8   | 0  | 1.640          | 1.641              | 1.670          |                |
| Nigeria              | 31 | 11 (3)  | 8  | 3  | 4.275 | 0.00338 | 4            | 0.546 | 2  | 1.766          | 1.440**            | 1.805**        |                |
| Venezuela            | 5  | 2 (2)   | 0  | 2  | 1.2   | 0.00095 | 2            | 0.6   | 0  | 1.459          | 1.459              | 1.432          |                |

All analyses were restricted to the signal peptide cleaved EtAMA1 ectodomain. Tajima's D and Fu and Li's D\* and F\* tests were used to assess of the extent or neutrality of signatures of selection with significance: \*\* $P < 0.05$ ; <sup>†</sup> $P < 0.02$ . All measures were calculated using DnaSP v5.10.01. dN, the number of nonsynonymous variant sites; dS, the number of synonymous variant sites; H, the number of sequence haplotypes detected; H<sup>†</sup>, the number of haplotypes specific to a region; Hd, the haplotype diversity; k, the average number of pairwise differences; N, the number of sequences tested;  $\pi$ , nucleotide diversity, calculated with the Jukes Cantor correction; S, the number of variant sites detected, with the number of parsimony-informative variant sites shown in parentheses.

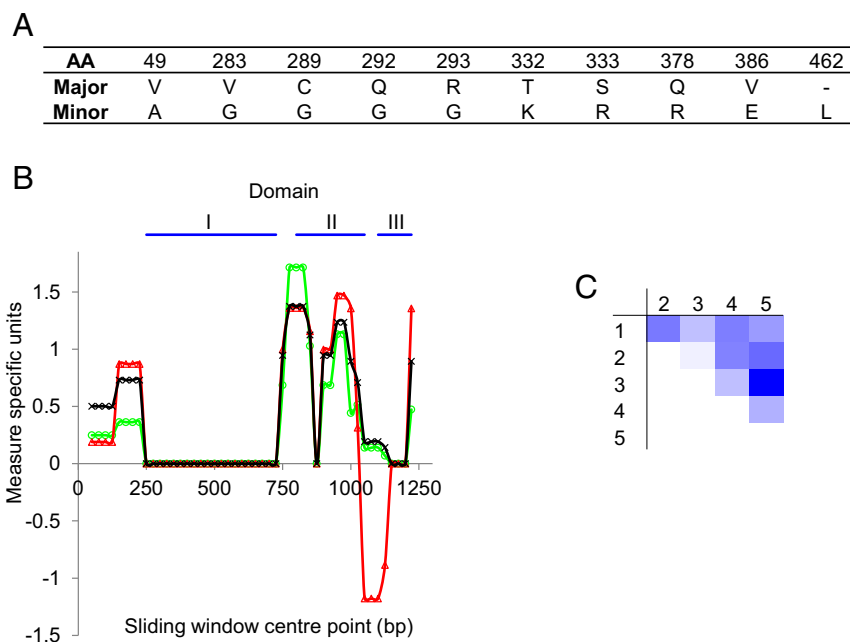
genome-wide polymorphism, only four of the eight *E. tenella ama1* haplotypes were region specific.

**Signatures of Selection Within *E. tenella* AMA1.** Antigens under deleterious immune selection commonly exhibit characteristic signatures of balancing selection. We applied Tajima's D and Fu and Li's D\* and F\* tests to the mature EtAMA1 ectodomain coding sequences to assess the occurrence and significance of any departure from neutral evolution. No significant evidence of balancing (or directional) selection was detected using Tajima's D across the full sequence or using a 100-bp sliding window with a 25-bp step size (Fig. 2B and Table 2). Both Fu and Li's D\* and F\* tests showed evidence of a low level of balancing selection acting primarily on domain II, although this finding was statistically significant only when assessed using the full sequence (Fig. 2B and Table 2). Codon-aligned nonsynonymous/synonymous substitutions (dN/dS) analysis of the mature ectodomain reinforced the overall view of low-level selective pressure with a ratio

of 0.823, indicating minimal purifying selection. To define global diversity, we calculated Wright's fixation index ( $F_{ST}$ ) for the full EtAMA1 dataset and each pairwise comparison between sequences pooled as Europe, Nigeria, Asia, the Americas, and North Africa (Fig. 2C). A global figure of 0.083 suggested a high level of parasite interbreeding, although pairwise analysis revealed greater genetic distance between regions, most notably between Asia and North Africa, highlighting the absence of parasite movement and/or strong balancing selection.

**Polyclonal Infection Is Common Within *E. tenella* Field Populations.**

The finding of high levels of region-specific genomic diversity in Nigeria and Southern India, but not in Northern India, Egypt and Libya, contrasted sharply with the limited diversity determined within the EtAMA1-coding sequence dataset. Returning to the Sequenom MassARRAY genotyping data, we found 46% of all samples submitted for typing to be polyclonal as determined by a multiple SNP identity at one or more loci. The ratio of polyclonal



**Fig. 2.** The extent and impact of sequence diversity within the *E. tenella* AMA11 (EtAMA1) coding sequence. (A) Nonsynonymous polymorphism within the EtAMA1 coding sequence. "Major" denotes the more common sequence type identified; "minor" the less common. (B) Sliding window plot of EtAMA1 coding sequence nucleotide diversity calculated using  $\pi$  with the Jukes Cantor correction (multiplied by 100; green line, open circle markers) and plots of neutrality using Tajima's D (red, open triangles) and Fu and Li's F\* (black, crosses). The window length was 100 bp with a 25-bp step size. No points were statistically significant. The EtAMA1 domain structure indicated by the blue lines was predicted based on cysteine residue locations. (C) Heat map representing the  $F_{ST}$  calculated for EtAMA1. Incremental intensity indicates elevated  $F_{ST}$ . 1, Europe; 2, Nigeria; 3, Asia; 4, the Americas; 5, North Africa.

to monoclonal populations was disproportionately higher in Southern India and Nigeria than in Northern India, Egypt, and Libya (72% and 64% compared with 10%, 12%, and 7%, respectively), likely as a consequence of the higher occurrence and larger number of haplotypes circulating. Nonetheless, opportunities for coinfection, and thus cross-fertilization, were detected in all regions sampled.

#### Cross-Fertilization Is Common During Polyclonal *E. tenella* Infection.

Polyclonal infection offers an opportunity for cross-fertilization but does not confirm its occurrence. Historic studies of cross-fertilization within *Eimeria* species have commonly used lethal drug or immune selection to enrich for hybrid progeny (e.g., refs. 14 and 25), obscuring the efficiency at which cross-fertilization occurs. In response to this knowledge gap, we undertook a series of coinfections between the Houghton and Nt2 *E. tenella* strains with oocyst ratios of 1:1, 2:1, 5:1, and 10:1 and a total dose of 500 sporulated oocysts per bird (two individual birds per ratio). Single sporocysts isolated from progeny oocysts were SNP-genotyped using Sequenom MassARRAY after FACS and whole-genome DNA amplification. In total, 36 sporocysts were genotyped for each dose ratio (Table 3). The proportion of hybrid sporocysts was highest for the 1:1 dose ratio and decreased as the ratio became more biased toward Houghton; nevertheless, hybrid parasites were detected within every sample. No single hybrid haplotype was detected more than twice, and the proportion of progeny oocysts arising from cross-fertilization always was higher than expected when assuming fertilization to have been a random event ( $P < 0.05$ ;  $\chi^2$  test).

#### Discussion

Understanding the population structure of *Eimeria* species parasites is of fundamental importance for explaining the epidemiology of coccidiosis, influencing the development of effective control strategies, and defining the genetic basis of control breakthrough. As recombinant or vectored subunit anticoccidial vaccines become technically more feasible, the regional occurrence of each *Eimeria* species and the extent of vaccine-targeted antigenic diversity assume greater relevance (5, 14). Under experimental conditions, studies with clonal reference strains of *Eimeria* have found antigens such as AMA1 to induce good levels of immunoprotection (14, 26). AMA1 also is an effective vaccine candidate in related apicomplexans including *Neospora caninum*, *T. gondii*, and several *Plasmodium* species (27–30), but its value for *Plasmodium* has been undermined by high levels of naturally occurring antigenic diversity (31, 32). In *Eimeria* the occurrence and extent of diversity at any antigen-encoding locus is completely unknown. Additionally, the frequency with which cross-fertilization occurs between genetically distinct *Eimeria* strains under field conditions is unclear, obscuring the relevance of meiotic segregation and recombination to the appearance of hybrid parasite genotypes capable of escaping chemoprophylaxis or vaccination. For novel subunit anticoccidial vaccines to proceed beyond the experimental phase, evidence of likely field efficacy and longevity in the face of immune selection is essential.

*Eimeria* genomes most commonly have been sampled by sequencing the 18S rDNA and ITS-1 and -2 sequences (e.g., ref. 33). ITS sequences often are used for molecular diagnostics and phylogenetic analyses because of the sensitivity and specificity provided by their high copy number and relatively rapid evolution, although intraclonal polymorphism has proven limiting (33, 34). Here, multiplex SNP genotyping for *E. tenella* provided far greater genetic power. Comparison of haplotype occurrence and diversity provided strong evidence of the expansion of a limited number of distinct haplotypes with significant LD in Northern India and North Africa, in contrast to Southern India and Nigeria, where no LD was detected. The limited haplotype diversity detected in the Northern regions partially resembles that described for the closely related coccidian *T. gondii*, for which a small number of dominant clonal genotypes have been described, supplemented by higher levels of genetic diversity in regions such as South America (35). However, comparison between these two genera reveals fundamental biological differences, including an apparently obligate requirement for sexual reproduction and a massive, regularly replaced, immunologically naive pool of potential definitive hosts (chickens) for *Eimeria*. In Southern India and Nigeria haplotypes were rarely identified more than once, suggesting that mating and recombination occur freely in these populations and that only a fraction of the genetic diversity may have been sampled. Such diversity has been interpreted previously in *Trypanosoma cruzi* as indicating high levels of parasite transmission, although limited clonal extinction also may be possible (36). Possible explanations for such variation include a higher *E. tenella* occurrence and elevated levels of poly-haplotype infection in regions found to harbor greater haplotype complexity, indicating more frequent opportunities for hybridization. The lower parasite occurrence in Northern India and North Africa would be expected to reduce opportunities for cross-fertilization, resulting in a more limited expansion of a smaller number of haplotypes with or without true clonal development. The causes of such clear differences in parasite occurrence are not yet apparent, although the greater poultry density in Southern as compared with Northern India does affect the availability of immunologically naive hosts, increasing opportunities for *Eimeria* replication (37). Larger distances between sampled farms in areas of greater haplotype diversity might have explained the variation detected, indicating greater isolation and opportunity for genetic drift. However, the farm distribution and sampling strategies were the same in Northern and Southern India and are unlikely to explain such distinct population structures. Those regions that were host to higher parasite occurrence and diversity also have been recognized as being more humid [subhumid to humid, compared with semiarid to hyper-arid (38)], potentially supporting higher rates of oocyst survival and sporulation and thus a greater environmental parasite burden.

Comparison of the cross-protective immunizing capacity of live *E. tenella* parasites from opposing ends of the NETWORK analysis, including the progenitor of the Paracox vaccinal line,

**Table 3. The impact of parental ratio during *E. tenella* coinfection on the likelihood of cross-fertilization (36 sporocysts per ratio)**

| Houghton:Nt2 ratio | No. birds | Sporocysts per bird | Self-fertilization         |                   |              | Cross-fertilization |                          |
|--------------------|-----------|---------------------|----------------------------|-------------------|--------------|---------------------|--------------------------|
|                    |           |                     | Expected (Houghton or Nt2) | Observed Houghton | Observed Nt2 | Expected            | Observed Houghton or Nt2 |
| 1:1                | 2         | 18                  | 18                         | 8                 | 5            | 18                  | 23                       |
| 2:1                | 2         | 18                  | 20                         | 11                | 7            | 16                  | 18                       |
| 5:1                | 2         | 18                  | 24                         | 20                | 2            | 12                  | 14                       |
| 10:1               | 2         | 18                  | 30                         | 24                | 2            | 6                   | 10                       |

Birds were coinfecting with Houghton and Nt2 *E. tenella* strains. Single-progeny sporocysts were identified as the product of self- or cross-fertilization following FACS purification and whole-genome amplification using Sequenom MassARRAY and comparison with both parental genotypes.

revealed significant strain-specific variation and some lack of heterologous immunoprotection, although the breakthrough in parasite replication was at a level unlikely to lead to occurrence of clinical disease (39). Antigenic diversity has been described elsewhere for *E. tenella* (40), albeit at a much lower level than reported for species such as *E. maxima* and *Eimeria acervulina* (14, 41, 42). The relevance of these findings to the other *Eimeria* species that infect poultry is not currently clear.

Previous genetic-mapping studies have shown AMA1 to be a target of protective anti-*Eimeria* immunity during secondary infection (14), indicating a likely balancing selective pressure, as has been described for loci encoding many other immunoprotective antigens (e.g., ref. 1). The nature of the immune response stimulated by AMA1 is not currently clear, although T cells and IFN- $\gamma$  are key components of the anti-*E. tenella* immune response (43), and anti-recombinant *E. tenella* AMA1 antibodies have been shown to inhibit the invasion of host cells by sporozoites (26). Analysis of *E. tenella* AMA1-coding sequences for signatures of selection revealed limited and conflicting evidence. Comparison of dN/dS ratios for the *E. tenella* AMA1-coding sequences suggested minimal to purifying selection, in agreement with previous cross-*Eimeria* species  $K_a/K_s$  analysis [ratio of the number of nonsynonymous substitutions per nonsynonymous site to the number of synonymous substitutions per synonymous site; gene ETH\_00007745,  $K_a/K_s = 0.12$  (13)], implying that conservation of AMA1 protein function may be of greater importance than immune escape. In contrast, application of Fu and Li's D\* and F\* tests indicated low levels of balancing selection. Combined, these findings suggest a low level of balancing selection within specific, possibly epitope-coding, regions that is consistent with the incomplete nature of the immune protection attributable to AMA1 revealed when tested as a recombinant protein or DNA vaccine (14, 26). The relatively short and apparently self-limiting *in vivo* *Eimeria* lifecycle also is likely to reduce the magnitude of antigen-specific immune selection. Prepatent periods define each *Eimeria* species with limited variation between strains, even when replicating in immunocompromised hosts such as MHC class I or II<sup>-/-</sup> knockout mice (11, 44). For *E. tenella*, microscopy-led studies have consistently identified three rounds of asexual replication before differentiation into gametocytes (45). The ability to select for serotinous lines characterized by extended prepatent periods, additional rounds of schizogony, and enhanced reproductive potential, apparently avoiding significant immune inhibition, indicates a key role for the parasite in the transition from asexual to sexual reproduction and ultimately termination of infection (46, 47). Parasite replication in the presence of minimal or incomplete immunity suggests that the host immune response has a weak role in the evolution of eimerian antigens such as AMA1, whose expression is restricted to the first zoite stage of the lifecycle (45). Alternatively, the relatively recent intensification of poultry production may indicate a similarly recent origin of intensified selection against antigens targeted by the host immune response, limiting any signature of selection. Competing selection incurred by routine chemoprophylaxis may reduce such signatures further. The low level of coding polymorphism detected here for AMA1 across temporally and spatially diverse parasite samples, combined with the cross-protection detected between genetically distinct parasite isolates when tested in outbred chickens, supports the validity of this antigen for use in future vaccine development. Additionally, these data suggest that for *Eimeria* the dN/dS ratio may not be an effective discriminator for identifying candidate vaccine antigens.

*Eimeria* are haploid for the majority of their life cycle with the exception of a short diploid phase during sexual reproduction. The appearance of genetic diversity can be accelerated by such a sexual phase, offering an opportunity for cross-fertilization and subsequent independent chromosomal segregation and recombination. The impact of genetic exchange during sexual replication on population structure is largely influenced by the balance between self- and cross-fertilization, which itself will vary by parasite availability as a

consequence of survival in the environment and chicken behavior (48). Thus, the frequency at which polyclonal infection occurs is a key determinant. Using a panel of 52 SNPs within an ~52 Mb genome (13), we found that 46% of all field samples containing *E. tenella* included at least two distinct genotypes. This finding is almost certainly an underestimate, and the inclusion of more markers or subjecting each sample to next-generation sequencing would likely increase this proportion. Quantitative modeling of SNP proportions at each locus could provide information about the clonal complexity of each sample (19), although this modeling would have required a greater number of replicates. Thus, although it is clear that opportunities for cross-fertilization would have been common for the parasites in our dataset, the proportion of cross-fertilization events per opportunity could not be determined. To shed light on the frequency of cross-fertilization events, we carried out single-sporocyst genotyping on the progeny of a series of experimental biconal infections and identified multiple hybrid haplotypes from every one of these infections. We used the sporocyst, which contains two identical haploid genomes, as the biological unit for this work to allow FACS isolation of single-clonal units after the completion of meiosis, reducing the risk of allelic drop out (49, 50). Comparison of the proportion of hybrid and parental genotypes revealed higher than expected numbers of hybrid progeny. In the absence of knowledge defining gametocyte ratios during coinfection with the two parasite lines used in the cross, a feature known to vary for the related apicomplexan *Plasmodium falciparum* (51), it was not possible to assess any bias between cross- and self-fertilization. Nonetheless, it is clear that cross-fertilization was common, highlighting the efficiency with which *E. tenella* can hybridize and indicating the ease with which vaccine- or drug-resistant alleles may be disseminated and mixed in field parasite populations. Corroboratory data from field parasite populations is scarce, but the large number of singleton haplotypes from Southern India and Nigeria is characteristic of populations with high rates of recombination (52).

Thus, we conclude that *Eimeria* species that infect chickens are ubiquitous and that at least one species, *E. tenella*, exhibits variable population structures with a concomitant variation in the probability of genetic exchange. Consequently, when resistance (e.g., to anticoccidial drugs or to possible future single antigen recombinant or subunit vaccines) emerges, it is likely to disseminate rapidly through production systems. However, eimerian antigens such as AMA1 retain promise as anticoccidial vaccine candidates, given the limited genetic and antigenic diversity observed and a conservative signature of selection. The incomplete nature of the protection induced by AMA1 immunization may prove beneficial to vaccine longevity by allowing local field strains to cycle at a level adequate to boost host immune protection and dilute genetic selection for vaccine resistance in a manner comparable to the success of ionophore-based chemoprophylaxis (53). Ultimately, combinations of two or more antigens are likely to be required to control each target *Eimeria* species.

## Materials and Methods

Detailed methods are included in *SI Materials and Methods*. All primers used were synthesized by Sigma-Genosys and are presented in [Table S2](#).

**Ethics Statement.** This study was carried out in strict accordance with the Animals (Scientific Procedures) Act 1986, an Act of Parliament of the United Kingdom. All animal studies and protocols were approved by the Royal Veterinary College Ethical Review Committee and the United Kingdom Government Home Office under the project licenses 30/2545 and 70/7781.

**Field Sample Selection.** A panel of 595 fecal samples was collected from eight countries ([Table S1](#)). India was most highly represented, with 288 samples collected from the states of Haryana, Punjab, Uttarakhand, and Uttar Pradesh in Northern India and from Andhra Pradesh, Karnataka, Kerala, and Tamil Nadu in Southern India ([Fig. S1](#)). Egypt and Libya were equally represented with 120 farms sampled for each, drawn from sites around Alexandria

and Cairo and around Benghazi, Zawilah, Misrata, Sirte, and Tripoli, respectively. Every sample was collected from a single pen per farm by walking a predetermined "W" pathway and collecting one fresh dropping every two to five paces (*SI Materials and Methods*). Thus, each sample represented multiple birds within a single cohort on a single farm. Because of differences in the availability of information, accessibility of farms, and legislation governing the exportation of biological samples, sampling frames using records from veterinary services, poultry suppliers, and farmer organizations were compiled using different approaches in each of the partner countries. Because the aim of the study was to sample for genetic diversity, farms were chosen to maximize parasite diversity, including intensive and extensive broiler and layer systems. Farms less than 100 m away from a previously sampled farm were excluded, as were farms with shared staff or equipment. In countries where 18 or more farms were sampled, those identified were compiled into lists of broiler and layer farms. A panel of farms representing each production type in each country then was randomly selected using Microsoft Excel randomizer (Microsoft Corporation). Samples were collected from one pen per farm where birds older than 3 wk of age were available.

**Parasites and Samples.** *E. tenella* isolates were maintained, purified, and propagated as described previously (11, 54, 55). Genomic DNA was isolated and purified using a QIAamp DNA Stool Mini kit (Qiagen) including an additional physical smashing step (56). Total RNA was prepared from sporulated oocyst field samples using the same approach, followed by purification using a Qiagen RNeasy mini kit as recommended by the manufacturer.

**Genetic Crosses: Parasites and Animals, Isolation of Single Sporocysts, and Genomic DNA Preparation.** The *E. tenella* Houghton and Nt2 strains were used for in vivo coinfection studies (13). Sporulated oocysts recovered from each genetic cross were disrupted mechanically to release sporocysts, which were sorted into 96-well plates using a FACSAria flow cytometer (Becton Dickinson); undertaken at the University of Oxford, Oxford, UK). Single sporocysts were disrupted by five cycles of freeze/thawing alternating between a methanol dry ice bath and the ambient temperature. Disrupted sporocysts were lysed and used as template for whole-genome amplification using a REPLI-g midi kit (Qiagen) as recommended by the manufacturer.

**Sequenom SNP Panel Design.** To ensure an even spread of polymorphisms throughout the *E. tenella* genome, a two-level approach was used in the design of the Sequenom assay SNP panel. The first level of selection focused on gene-specific SNPs. Sequence polymorphisms in chromosomally assigned

genes and genes encoding proteins that function at the host/parasite interface were compared for the *E. tenella* Houghton and Nt2 assemblies (*Dataset S1*) (13, 17). Gene-specific polymorphisms were selected manually using the local BLAST and sequence comparison tools in CLC Main Workbench (CLC Bio). *E. tenella* genome and HAPPY (mapping based on the analysis of approximately HAPloid DNA samples using Polymerase chain reaction) mapping information was used where possible to attribute each gene-specific SNP to a chromosome to give an even dispersal of gene-specific SNPs across the 14 *E. tenella* chromosomes (13, 17).

The second level of SNP selection was based on intrachromosomal polymorphism in chromosomes 1 and 2. The R and P segmental structure of chromosome 1 and 2 (13, 18) was used as a template for SNP selection by placing a minimum of two SNPs in each R and P region (*Dataset S1*). R and P regions were designated from published parameters for chromosome 1 (18) and from contigs annotated as R and P regions for chromosome 2 (13). The Genome Analysis Toolkit (GATK, version 2.6-4; Broad Institute) was used to call SNPs for the nucleotides against the Houghton chromosome 1 and 2 assemblies and then to filter for quality (57, 58). SNPs were selected manually by visual comparison in Artemis (version 15; Sanger Institute) for quality across multiple sequence reads and positions within the chromosome to give a uniform distribution within each region (59).

Oligonucleotides for the Sequenom MassARRAY were designed by Sequenom within the left and right 150-bp SNP-flanking regions. Flanking regions for the panel of SNPs were identified manually based on the *E. tenella* Houghton assembly in CLC Main Workbench and chromosome 1 and 2 using in-house scripts. The MassARRAY genotyping assays were optimized for multiplexing by Sequenom, which separated the panel of SNPs into two separate multiplexes comprising 55 SNPs in total (*Dataset S1*) (60).

**Genetics Analysis.** Software used in these studies to analyze the MassARRAY-generated SNPs included DnaSP v5.10.01 (20), NETWORK v4.6.1.1 (22), STRUCTURE v2.3.4 (21), and LIAN v3.6 (23). Gene-specific population analysis used DnaSP v5.10.01 and SNAP v2.1.1 (61).

**ACKNOWLEDGMENTS.** We thank S. Kumar, S. Aarthi, R. Selvabharathi, Dr. Abdul S. Chaudhry, Prof. Olivier Sparagano, Prof. Idris A. Lawal, Prof. J. K. P. Kwaga, and Prof. A. J. Nok for assistance in arranging for the collection of samples used in these studies. The work described here was funded by the UK Department for International Development and the Biotechnology and Biological Sciences Research Council through the Combating Infectious Diseases of Livestock for International Development initiative under Project Reference BB/H009337. This paper has been assigned the reference no. PPB\_00975 by the Royal Veterinary College.

- Amambua-Ngwa A, et al. (2012) Population genomic scan for candidate signatures of balancing selection to guide antigen characterization in malaria parasites. *PLoS Genet* 8(11):e1002992.
- Minot S, et al. (2012) Admixture and recombination among *Toxoplasma gondii* lineages explain global genome diversity. *Proc Natl Acad Sci USA* 109(33):13458–13463.
- Manske M, et al. (2012) Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. *Nature* 487(7407):375–379.
- Beck HP, et al. (2009) Molecular approaches to diversity of populations of apicomplexan parasites. *Int J Parasitol* 39(2):175–189.
- Blake DP, Tomley FM (2014) Securing poultry production from the ever-present *Eimeria* challenge. *Trends Parasitol* 30(1):12–19.
- O'Neill BC, et al. (2010) Global demographic trends and future carbon emissions. *Proc Natl Acad Sci USA* 107(41):17521–17526.
- Smil V (2002) Worldwide transformation of diets, burdens of meat production and opportunities for novel food proteins. *Enzyme Microb Technol* 30(3):305–311.
- Godfray HC, et al. (2010) Food security: The challenge of feeding 9 billion people. *Science* 327(5967):812–818.
- Perry B, Randolph T, McDermott J, Sones K, Thornton P (2002) *Investing in Animal Health Research to Alleviate Poverty* (International Livestock Research Institute), Nairobi, Kenya.
- Dalloul RA, Lillehoj HS (2006) Poultry coccidiosis: Recent advancements in control measures and vaccine development. *Expert Rev Vaccines* 5(1):143–163.
- Shirley MW, Smith AL, Tomley FM (2005) The biology of avian *Eimeria* with an emphasis on their control by vaccination. *Adv Parasitol* 60:285–330.
- Peek HW, Landman WJ (2011) Coccidiosis in poultry: Anticoccidial products, vaccines and other prevention strategies. *Vet Q* 31(3):143–161.
- Reid AJ, et al. (2014) Genomic analysis of the causative agents of coccidiosis in domestic chickens. *Genome Res* 24(10):1676–1685.
- Blake DP, et al. (2011) Genetic mapping identifies novel highly protective antigens for an apicomplexan parasite. *PLoS Pathog* 7(2):e1001279.
- Chapman HD, Shirley MW (2003) The Houghton strain of *Eimeria tenella*: A review of the type strain selected for genome sequencing. *Avian Pathol* 32(2):115–127.
- Blake DP, et al. (2004) Parasite genetics and the immune host: Recombination between antigenic types of *Eimeria maxima* as an entrée to the identification of protective antigens. *Mol Biochem Parasitol* 138(1):143–152.
- Lim LS, Tay YL, Alias H, Wan KL, Dear PH (2012) Insights into the genome structure and copy-number variation of *Eimeria tenella*. *BMC Genomics* 13(1):389.
- Ling KH, et al. (2007) Sequencing and analysis of chromosome 1 of *Eimeria tenella* reveals a unique segmental organization. *Genome Res* 17(3):311–319.
- Schmid-Hempel P, Reber Funk C (2004) The distribution of genotypes of the trypanosome parasite, *Crithidia bombi*, in populations of its host, *Bombus terrestris*. *Parasitology* 129(Pt 2):147–158.
- Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25(11):1451–1452.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155(2):945–959.
- Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16(1):37–48.
- Haubold B, Hudson RR (2000) LIAN 3.0: Detecting linkage disequilibrium in multilocus data. *Linkage Analysis. Bioinformatics* 16(9):847–848.
- Kundu K, et al. (2014) Cloning and sequencing of beta-tubulin and internal transcribed spacer-2 (ITS-2) of *Eimeria tenella* isolate from India. *J Parasit Dis*, 10.1007/s12639-013-0392-4.
- Joyner LP, Norton CC (1975) Transferred drug-resistance in *Eimeria maxima*. *Parasitology* 71(3):385–392.
- Jiang L, et al. (2012) Identification and characterization of *Eimeria tenella* apical membrane antigen-1 (AMA1). *PLoS One* 7(7):e41115.
- Zhang H, et al. (2010) Neospora caninum: Application of apical membrane antigen 1 encapsulated in the oligomannose-coated liposomes for reduction of offspring mortality from infection in BALB/c mice. *Exp Parasitol* 125(2):130–136.
- Hodder AN, Crewther PE, Anders RF (2001) Specificity of the protective antibody response to apical membrane antigen 1. *Infect Immun* 69(5):3286–3294.
- Crewther PE, Matthew ML, Flegg RH, Anders RF (1996) Protective immune responses to apical membrane antigen 1 of *Plasmodium chabaudi* involve recognition of strain-specific epitopes. *Infect Immun* 64(8):3310–3317.
- Dautu G, et al. (2007) *Toxoplasma gondii*: DNA vaccination with genes encoding antigens MIC2, MZAP, AMA1 and BAG1 and evaluation of their immunogenic potential. *Exp Parasitol* 116(3):273–282.

31. Eisen DP, Saul A, Fryauff DJ, Reeder JC, Coppel RL (2002) Alterations in *Plasmodium falciparum* genotypes during sequential infections suggest the presence of strain specific immunity. *Am J Trop Med Hyg* 67(1):8–16.
32. Healer J, et al. (2004) Allelic polymorphisms in apical membrane antigen-1 are responsible for evasion of antibody-mediated inhibition in *Plasmodium falciparum*. *Mol Microbiol* 52(1):159–168.
33. 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. *J Parasitol* 95(4):871–880.
34. Schnitzler BE, Thebo PL, Tomley FM, Uggla A, Shirley MW (1999) PCR identification of chicken *Eimeria*: A simplified read-out. *Avian Pathol* 28(1):89–93.
35. Su C, et al. (2012) Globally diverse *Toxoplasma gondii* isolates comprise six major clades originating from a small number of distinct ancestral lineages. *Proc Natl Acad Sci USA* 109(15):5844–5849.
36. Ramírez JD, et al. (2012) Contemporary cryptic sexuality in *Trypanosoma cruzi*. *Mol Ecol* 21(17):4216–4226.
37. USDA (2013) Poultry in India. *International Egg and Poultry Review* 16(46):1–3.
38. Deichmann U, Eklundh L (1991) *Global Digital Data Sets for Land Degradation Studies: A GIS Approach* (United Nations Environment Program/Global Environment Monitoring System and Global and Regional Integrated Data Centers, Nairobi, Kenya).
39. McDonald V, Shirley MW, Millard BJ (1986) A comparative study of two lines of *Eimeria tenella* attenuated either by selection for precocious development in the chicken or by growth in chicken embryos. *Avian Pathol* 15(3):323–335.
40. Awad AM, El-Nahas AF, Abu-Akkada SS (2013) Evaluation of the protective efficacy of the anticoccidial vaccine Coccivac-B in broilers, when challenged with Egyptian field isolates of *E. tenella*. *Parasitol Res* 112(1):113–121.
41. Joyner LP (1969) Immunological variation between two strains of *Eimeria acervulina*. *Parasitology* 59(3):725–732.
42. Blake DP, et al. (2012) EmaxDB: Availability of a first draft genome sequence for the apicomplexan *Eimeria maxima*. *Mol Biochem Parasitol* 184(1):48–51.
43. Hong YH, Lillehoj HS, Lee SH, Dalloul RA, Lillehoj EP (2006) Analysis of chicken cytokine and chemokine gene expression following *Eimeria acervulina* and *Eimeria tenella* infections. *Vet Immunol Immunopathol* 114(3–4):209–223.
44. Smith AL, Hayday AC (2000) Genetic dissection of primary and secondary responses to a widespread natural pathogen of the gut, *Eimeria vermiformis*. *Infect Immun* 68(11):6273–6280.
45. Lal K, et al. (2009) Proteomic comparison of four *Eimeria tenella* life-cycle stages: Unsporulated oocyst, sporulated oocyst, sporozoite and second-generation merozoite. *Proteomics* 9(19):4566–4576.
46. Williams RB (2007) “Serotinous”, a descriptor proposed for lines of *Eimeria* species (Apicomplexa: Eimeriidae) selected for delayed development: An antonym for “precocious”. *Vet Parasitol* 145(3–4):388–389.
47. Dong H, Suo X, Wang M, Teng K (2006) Characteristics of a line of *Eimeria necatrix* after 16 successive passages of oocysts collected after peak oocyst production. *J Parasitol* 92(6):1229–1234.
48. Srinivasa Rao AS, Tomley F, Blake D (2014) Understanding chicken walks on n x n grid: Hamiltonian paths, discrete dynamics, and rectifiable paths. *Math Methods Appl Sci*, 10.1002/mma.3301.
49. Shirley MW, Harvey DA (1996) *Eimeria tenella*: Infection with a single sporocyst gives a clonal population. *Parasitology* 112(Pt 6):523–528.
50. Nair S, et al. (2014) Single-cell genomics for dissection of complex malaria infections. *Genome Res* 24(6):1028–1038.
51. Reece SE, Drew DR, Gardner A (2008) Sex ratio adjustment and kin discrimination in malaria parasites. *Nature* 453(7195):609–614.
52. Yeo M, et al. (2011) Multilocus sequence typing (MLST) for lineage assignment and high resolution diversity studies in *Trypanosoma cruzi*. *PLoS Negl Trop Dis* 5(6):e1049.
53. Chapman HD (1976) *Eimeria tenella* in chickens: Studies on resistance to the anticoccidial drugs monensin and lasalocid. *Vet Parasitol* 2(2):187–196.
54. Long PL, Millard BJ, Joyner LP, Norton CC (1976) A guide to laboratory techniques used in the study and diagnosis of avian coccidiosis. *Folia Vet Lat* 6(3):201–217.
55. Ministry of Agriculture, Fisheries and Food (1986) *Manual of Veterinary Parasitological Laboratory Techniques* (Her Majesty's Stationary Office, London).
56. Kumar S, et al. (2014) An optimised protocol for molecular identification of *Eimeria* from chickens. *Vet Parasitol* 199(1–2):24–31.
57. McKenna A, et al. (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20(9):1297–1303.
58. DePristo MA, et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43(5):491–498.
59. Rutherford K, et al. (2000) Artemis: Sequence visualization and annotation. *Bioinformatics* 16(10):944–945.
60. Storm N, Darnhofer-Patel B, van den Boom D, Rodi CP (2003) MALDI-TOF mass spectrometry-based SNP genotyping. *Methods Mol Biol* 212:241–262.
61. Korber B (2000) HIV Signature and Sequence Variation Analysis. *Computational Analysis of HIV Molecular Sequences*, eds Rodrigo A, Learn G (Kluwer Academic Publishers, Dordrecht, The Netherlands), pp 55–72.
62. Fornace KM, et al. (2013) Occurrence of *Eimeria* species parasites on small-scale commercial chicken farms in Africa and indication of economic profitability. *PLoS One* 8(12):e84254.
63. Vrba V, Blake DP, Poplstein M (2010) Quantitative real-time PCR assays for detection and quantification of all seven *Eimeria* species that infect the chicken. *Vet Parasitol* 174(3–4):183–190.
64. 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. *Int J Parasitol* 36(1):97–105.
65. Nolan MJ, Tomley FM, Kaiser P, Blake DP (2015) Quantitative real-time PCR (qPCR) for *Eimeria tenella* replication - Implications for experimental refinement and animal welfare. *Parasitol Int* 64(5):464–470.
66. Eckert J, Braun R, Shirley M, Coudert P (1995) *Guidelines on Techniques In Coccidiosis Research* (European Commission, Brussels).
67. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol Ecol* 14(8):2611–2620.
68. Jukes T, Cantor C (1969) Evolution of protein molecules. *Mammalian Protein Metabolism*, ed Munro H (Academic, New York), pp 21–132.
69. Nei M (1987) *Molecular Evolutionary Genetics* (Columbia Univ Press, New York).
70. Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123(3):585–595.
71. Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics* 133(3):693–709.
72. Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. *Genetics* 132(2):583–589.