

# The apicoplast genome of *Leucocytozoon caulleryi*, a pathogenic apicomplexan parasite of the chicken

Takayuki Imura · Shigeharu Sato · Yukita Sato ·  
Daichi Sakamoto · Takashi Isobe · Koichi Murata ·  
Anthony A. Holder · Masayoshi Yukawa

Received: 16 August 2013 / Accepted: 18 November 2013 / Published online: 4 December 2013  
© The Author(s) 2013. This article is published with open access at Springerlink.com

**Abstract** *Leucocytozoon caulleryi*, a haemosporidian parasite of the chicken (*Gallus gallus domesticus*), can be highly pathogenic and often fatal. Although this parasite is extremely relevant to veterinary science, knowledge of its genomic features is limited. To gain information applicable to developing novel control methods for the parasite, we analyzed the apicoplast genome of *L. caulleryi*. This extranuclear organellar DNA of 85.1 % A+T and a unit of 34,779 bp was found to encode almost the same set of genes as the plastid genome of *Plasmodium falciparum*, including 16 tRNA and 30 protein coding genes, and except for one open reading frame, *ORF91* absent in *L. caulleryi*. As in *P. falciparum*, the *L. caulleryi* apicoplast DNA contains two sets of a unique inverted repeat (IR), each one 5,253 bp and encoding genes specifying one large and one small rRNA subunit and nine tRNAs but no protein, and separated by a unique 13 bp sequence. Studies of several haemosporidian

apicoplast DNA sequences have identified a corresponding IR region; however, none of these studies has looked at the complete sequence, even for well-studied species such as *P. falciparum*. Phylogenetic studies using a concatenated amino acid sequence based on the open reading frames confirmed the close relationship between *L. caulleryi* and *Plasmodium* spp. In this study, we determined the nucleotide sequence of the entire *L. caulleryi* apicoplast genome, including the region connecting the two IR units. This is the first report of the complete nucleotide sequence of a haemosporidian apicoplast DNA with a canonical IR.

## Introduction

The phylum Apicomplexa consists of protistan parasites, including species of medical or veterinary importance such as *Plasmodium falciparum*, *Toxoplasma gondii*, and *Eimeria tenella*. One of the characteristics shared by these parasites is the presence of a vestigial plastid called the apicoplast (Foth and McFadden 2003; McFadden 2011). This plastid has been suggested to have originated from an ancient endosymbiotic alga of debated phylogenetic position (Köhler et al. 1997; Janouskovec et al. 2010). Previous studies have suggested that this organelle is involved in critical metabolic processes such as heme and isoprenoid biosynthesis (McFadden 2011; Ralph et al. 2004). Indeed, *P. falciparum* that has lost its apicoplast cannot grow unless the product of the apicoplastic isoprenoid biosynthesis, isopentenyl pyrophosphate, is supplemented to the culture medium (Yeh and DeRisi 2011). These studies have led to recognition of the apicoplast as a potential anti-*Plasmodium* drug target.

As a vestigial plastid, the apicoplast retains its own genome. To date, analysis of the apicoplast genomes of nine *Plasmodium* species, *T. gondii*, *E. tenella*, *Theileria parva*, and *Babesia bovis* has been performed (Arisue et al. 2012;

Takayuki Imura and Shigeharu Sato contributed equally to this study.

T. Imura · Y. Sato · D. Sakamoto · M. Yukawa  
Department of Veterinary Medicine, College of Bioresource  
Sciences, Nihon University, Fujisawa, Kanagawa 252-0880, Japan

S. Sato · A. A. Holder  
Division of Parasitology, MRC National Institute for Medical  
Research, London NW7 1AA, UK

T. Isobe  
National Institute of Animal Health, Tsukuba, Ibaraki 305-0856,  
Japan

K. Murata  
Department of Animal Resource Sciences, College of Bioresource  
Sciences, Nihon University, Fujisawa, Kanagawa 252-0880, Japan

Y. Sato (✉)  
Laboratory of Biomedical Science, Department of Veterinary  
Medicine, Nihon University, Fujisawa, Kanagawa 252-0880, Japan  
e-mail: sato.yukita@nihon-u.ac.jp

Brayton et al. 2007; Cai et al. 2003; Gardner et al 2005; Lau et al. 2009; Williamson et al. 2001), and the results recorded in gene databases. These data revealed that features of the apicoplast genome are fairly well conserved between species, although they may also differ. The part of the apicoplast DNA sequence encoding the 2 rRNA genes (*rrl*, *rrs*) as well as genes for 9 tRNA species (*trnA*(UGC), *trnI*(GAU), *trnL*(UAG), *trnM*(CAU), *trnN*(GUU), *trnR*(ACG), *trnR*(UCU), *trnT*(UGU), and *trnV*(UAC)) are duplicated to form an inverted repeat (IR) in all *Plasmodium* species studied, except for the rodent malaria parasite *Plasmodium chabaudi chabaudi* (Arisue et al. 2012; Sato et al. 2013). An IR with identical features exists even in coccidian apicoplast sequences, although the corresponding sequence is not duplicated in piroplasmids. These findings suggest that the

apicoplast IR was present in the common ancestor of all extant apicomplexan species and that one copy of the duplicated IR units was independently lost in the branches corresponding to piroplasmids and *P. chabaudi chabaudi* (Sato et al. 2013).

The apicomplexan genus *Leucocytozoon* consists of species which parasitize birds (Valkiunas 2005). Among them, *L. caulleryi*, which can cause fatal malaria-like symptoms in the chicken (*Gallus gallus domesticus*), is one of the species with the greatest economic impact, and it is distributed mainly throughout Asian countries including Japan (Morii et al. 1981; Morii 1992; Yu and Wang 2001). Genomic information could contribute not only to understanding the biological features of pathogens but also to the development of control procedures for infection, such as genetically recombinant vaccines or protein-targeted drugs. Trials to induce production

**Table 1** Comparison of apicoplast genomes

Species	<i>Leucocytozoon caulleryi</i> Niigata	<i>Plasmodium falciparum</i> C10	<i>Plasmodium chabaudi chabaudi</i> CB	<i>Toxoplasma gondii</i> RH	<i>Eimeria tenella</i> Penn State	<i>Theileria parva</i> Muguga	<i>Babesia bovis</i> T2Bo
Complete/partial	Complete	Partial	Complete	Complete	Complete	Complete	Complete
Size (bp)	34,779	34,682	29,623	34,996	34,750	39,579	33,351
A+T(%)	85.1	86.9	86.3	78.6	79.4	80.5	78.2
IR <sup>z</sup>	+ <sup>a</sup>	+	–	+	+	–	–
Protein <sup>b,c</sup>	30	31	31	29	29	32 <sup>d</sup>	30 <sup>e</sup>
<i>clpC</i> <sup>b</sup>	1	1	1	1	1	2	2
<i>sufB</i>	+	+	+	+	+	–	–
<i>rpl23</i>	+	+	+	–	–	–	–
<i>ORF91</i>	–	+	+	–	–	–	–
rRNA <sup>b</sup>	2 <sup>f</sup>	2 <sup>f</sup>	2	2 <sup>f</sup>	2 <sup>f</sup>	2	2
tRNA <sup>b</sup>	25 <sup>g</sup>	25 <sup>g</sup>	26 <sup>h</sup>	24 <sup>g</sup>	24 <sup>g</sup>	24 <sup>i</sup>	24
<i>trnG</i> (ACC)	+	+	+	–	–	–	–
Intron in <i>trnL</i> (UUA)	+	+	+	+	–	+	+
In-frame stop codon <sup>j</sup>	–	–	–	+	+	–	–
Coding strand	Both	Both	Both	Both	Both	One	One
Accession number	AP013071	X95275, X95276	HF563595	U87145	AY217738	NC_007758	NC_011395

<sup>z</sup> Inverted repeat

<sup>a</sup> + present, – absent

<sup>b</sup> Number of different products encoded

<sup>c</sup> *rpoC2* considered two separate proteins (*rpoC2A* and *rpoC2B*)

<sup>d</sup> Excluding those specified by 12 repetitive genes predicted between *rpoC1* and *rpoC2A*

<sup>e</sup> Excluding repetitive gene products from BBOV\_V000300, BBOV\_V000310, BBOV\_V000320, BBOV\_V000180, BBOV\_V000200, and BBOV\_V000210

<sup>f</sup> The gene for each is present in IR and duplicated

<sup>g</sup> Genes specifying 9 tRNA species (tRNA-Ala(UGC), Arg(ACG), Arg(UCU), Asn(GUU), Ile(GAU), Leu(UAG), Met(CAU), Thr(UGU), and Val(UAC)) are present in IR and duplicated

<sup>h</sup> Two genes encode different versions of tRNA-Thr. The gene for tRNA-Met is duplicated

<sup>i</sup> Five genes encode the same tRNA-Met(CAU)

<sup>j</sup> Excluding the hypothetical codon connecting *rpoC2A* and *rpoC2B*



the nucleotide sequence of each PCR fragment was determined directly by Sanger sequencing as described previously (Omori et al. 2008).

### High throughput sequence analysis of genomic DNA

A library for high-throughput sequencing (HTS) was prepared from the parasite DNA using the Paired End Sample Prep Kit (Illumina) and analyzed by the Illumina method on a Genome Analyzer Iix (Illumina). From the HTS data obtained, reads matching the *trnI* gene sequence at the end of the IR unit were searched for by using BLASTN to query the NCBI database. Obtained reads were aligned using Clustal X (Larkin et al. 2007) to extend the nucleotide sequences in locations where the PCR direct sequencing method described above was insufficient.

The position and the anticodon of each tRNA gene encoded by the *L. caulleryi* plastid DNA sequence were predicted by the tRNAscan-SE server (<http://lowelab.ucsc.edu/tRNAscan-SE/>). All other genes were manually annotated based on homology to *P. falciparum*.

### Phylogenetic analysis

Phylogeny of the plastid genomes of *L. caulleryi* and 7 other apicomplexan species (*P. falciparum*, *P. gallinaceum*, *P. berghei*, *T. parva*, *B. bovis*, *E. tenella*, and *T. gondii*) as well as *Chromera velia*, which was chosen as the out-group, was deduced from the amino acid sequences that represent the genomes. The representative sequences, which were 2,478 residues long, were created by concatenating the amino acid sequences of the products of the following 10 genes: *rpl 2*, *rpl 6*, *rpl 14*, *rpl 16*, *rps 2*, *rps 3*, *rps 11*, *rps 12*, *tufA*, and *rpoB*. An alignment of these sequences was created using Clustal W (Larkin et al. 2007), and the phylogeny of these sequences was analyzed using MEGA5 (<http://www.megasoftware.net>) by the maximum likelihood method with the JTT+F+Γ model. Bootstrap analysis with 1,000 replicates was performed to estimate the confidence of the tree topology.

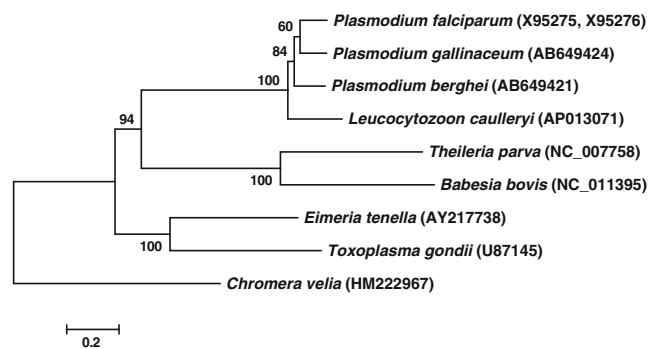
## Results

We determined the complete nucleotide sequence of the *Leucocytozoon* plastid DNA for the first time. The *L. caulleryi* plastid DNA was 34,779 bp, and its A+T content was 85.1 % (Table 1). These values were within the range of other plastid DNA data previously reported. Like the plastid DNA of *Plasmodium* spp. and the coccidians, the plastid DNA of *L. caulleryi* had an inverted repeat (IR) that occupied about 1/3 of the unit sequence (Fig. 1). We had no problem in amplifying the majority of the plastid DNA by PCR. However, the sequence connecting the two IR units at the end closest to the tRNA-Ile gene was the exception; we had no success in

obtaining PCR fragments that contained the short sequence (data not shown). It is possible that this region of the *L. caulleryi* plastid DNA forms a peculiar structure that interferes with extension of the DNA strand in PCR, similar to the “tip region” present between the two IR units of the *E. tenella* plastid DNA (Cai et al. 2003). Therefore, we analyzed the total genomic DNA extracted from the parasite preparation by whole genome sequencing with the Illumina platform, and we searched for reads corresponding to the end of the IR unit in these HTS data. These identified reads filled the gap between the two IR units, which was found to be only 13 bp, allowing us to finally obtain the complete plastid nucleotide sequence.

Each IR unit, which was 5,253 bp, encoded one gene each of large and small ribosomal RNA subunits (*rrl* and *rrs*, respectively) as well as 9 tRNA genes (*trnA*(UGC), *trnI*(GAU), *trnL*(UAG), *trnM*(CAU), *trnN*(GUU), *trnR*(ACG), *trnR*(UCU), *trnT*(UGU), and *trnV*(UAC)), but no protein-coding gene. These characteristics are conserved across all presently studied apicomplexan species with IR-positive apicoplast DNA.

There were 30 total protein-coding genes annotated in the *L. caulleryi* plastid DNA sequence. These genes, as well as 16 tRNA genes, were encoded outside the IR region. Both the content of the genome itself and the arrangement of each gene on the DNA were remarkably highly conserved between *L. caulleryi* and *P. falciparum*. The only difference between the two was the presence of the *ORF91* gene, which is missing from the *L. caulleryi* genome but encoded between *rps5* and *rpl36* (Wilson et al. 1996) in *P. falciparum* DNA. Interestingly, this open reading frame (ORF) is also absent from coccidian and piroplasmid plastid genomes (Table 1), despite the fact that *ORF91* in each *Plasmodium* species specifies a fairly conserved amino acid sequence (data not shown).



**Fig. 2** Phylogeny of the apicoplast genomes. Amino acid sequences specified by the 10 selected plastid-encoded genes were concatenated for each species, and the phylogeny was analyzed by the maximum likelihood method with the JTT+F+Γ model to obtain the dendrogram. Numbers at nodes indicate bootstrap values from 1,000 replicates. The scale bar indicates 0.2 substitutions per site. The accession numbers of the nucleotide sequence data, from which the amino acid sequences used in the analysis were deduced, are given in parentheses after the name of each taxon

Phylogenetic analysis of the concatenated amino acid sequences of 10 gene products (2,478 residues) suggested a close relationship between *L. caulleryi* and *Plasmodium* (Fig. 2). This finding, along with similarities in both gene content and arrangement of genes on the DNA, suggests that the *L. caulleryi* plastid has similar functions to that of *Plasmodium*.

## Discussion

By analogy to *P. falciparum* DNA (Williamson et al. 2002), we had speculated that most of the *L. caulleryi* plastid DNA exists as circular DNA rather than linear multimers, as in *T. gondii* (Williamson et al. 2001). This organization, together with the presence of a set of inverted repeats was confirmed by the sequence analysis.

Phylogenetic studies using apicoplast genome sequences have been used to examine relationships between members of the Apicomplexa. For example, Oborník and colleagues (2002) used 894 nt apicoplast DNA sequences corresponding to part of the *rrs* gene and showed this provide a useful tool to analyze Apicomplexa and distinguish between coccidian and haemosporidian parasites. Our use of concatenated amino acid sequences extends this analysis and suggests a close relationship between *L. caulleryi* and *Plasmodium*.

Both the organization of the genes and the phylogenetic analysis of amino acid sequences suggested that the *L. caulleryi* plastid DNA is very similar to that of *Plasmodium* spp. One notable difference is the absence of *ORF91* from the *L. caulleryi* plastid genome. It is as yet unknown whether the missing gene has been translocated to the nuclear genome so that its protein product is imported into the plastid; however, the absence of *ORF91* in the organellar genome suggests that, unlike *Plasmodium* spp., non-*Plasmodium* species lack the plastid function in which the *ORF91* gene product is involved.

As with other apicoplast DNA, the *rpoC2* gene in *L. caulleryi* specifying a specific subunit of the multi subunit-type DNA-dependent RNA polymerase was interrupted by an in-frame UGA codon. Here we annotated the two juxtaposing ORFs separately as *rpoC2A* and *rpoC2B*, although these might be read through in translation to produce a continuous polypeptide as hypothesized for *rpoC2* in *P. falciparum* (Wilson et al. 1996). Unlike coccidians (<http://roos.bio.upenn.edu/~rooslab/jkissing/toxomap.html>, Cai et al. 2003), no other apparent in-frame stop codon has been predicted in the plastid genome of *L. caulleryi*.

The plastid, which is known to provide the platform for critical metabolic processes such as isoprenoid biosynthesis, has been recognized as a promising target for drugs to inhibit the growth of *Plasmodium* spp., and understanding its metabolic functions may aid in designing drugs that prevent the

infection and transmission of *L. caulleryi*. Elucidation of its genome structure is an important step along this pathway.

The apicoplast genome of *L. caulleryi* is much larger than the mitochondrial genome of this species (Omori et al. 2008). In the absence of the apicoplast sequence, phylogenetic analysis and genetic typing of *Leucocytozoon* had been generally carried out using mitochondrial sequence such as that of the cytochrome *b* gene alone (Imura et al. 2010; Ortego and Cordero 2009). Now we have established the complete nucleotide sequence of the *Leucocytozoon* apicoplast genome, this newly available information will contribute to improving the accuracy of both phylogenetic analysis and genetic typing of apicomplexan parasites including *Leucocytozoon*.

**Acknowledgments** This study was supported in part by a Grant-in-Aid for Scientific Research (No. 21580406) from the Japan Society for the Promotion of Science and the Strategic Research Base Development Program “International research on epidemiology of zoonoses and training for young researchers” matching fund subsidy from the Ministry of Education, Culture, Sports, Science and Technology of Japan. Work in the Holder laboratory was supported by the UK Medical Research Council (File References U117532067 and U117584270).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Arisue N, Hashimoto T, Mitsui H, Palacpac NM, Kaneko A, Kawai S, Hasegawa M, Tanabe K, Horii T (2012) The Plasmodium apicoplast genome: conserved structure and close relationship of *P. ovale* to rodent malaria parasites. *Mol Biol Evol* 29:2095–2099
- Brayton KA, Lau AO, Herndon DR, Hannick L, Kappmeyer LS, Berens SJ, Bidwell SL, Brown WC, Crabtree J, Fadrosch D, Feldblum T, Forberger HA, Haas BJ, Howell JM, Khouri H, Koo H, Mann DJ, Norimine J, Paulsen IT, Radune D, Ren Q, Smith RK Jr, Suarez CE, White O, Wortman JR, Knowles DP Jr, McElwain TF, Nene VM (2007) Genome sequence of *Babesia bovis* and comparative analysis of apicomplexan haemoprotozoa. *PLoS Pathol* 3:1401–1413
- Cai X, Fuller AL, McDougald LR, Zhu G (2003) Apicoplast genome of the coccidian *Eimeria tenella*. *Gene* 321:39–46
- Foth BJ, McFadden GI (2003) The apicoplast: a plastid in *Plasmodium falciparum* and other Apicomplexan parasites. *Int Rev Cytol* 224: 57–110
- Gardner MJ, Bishop R, Shah T, de Villiers EP, Carlton JM, Hall N, Ren Q, Paulsen IT, Pain A, Berriman M, Wilson RJM, Sato S, Ralph SA, Mann DJ, Xiong Z, Shallom SJ, Weidman J, Jiang L, Lynn J, Weaver B, Shoaibi A, Domingo AR, Wasowo D, Crabtree J, Wortman JR, Haas B, Angiuoli SV, Creasy TH, Lu C, Suh B, Silva JC, Utterback TR, Feldblum TV, Pertea M, Allen J, Nierman WC, Taracha ELN, Salzberg SL, White OR, Fitzhugh HA, Morzaria S, Venter JC, Fraser CM, Nene V (2005) Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. *Science* 309:134–137
- Imura T, Sato Y, Ejiri H, Tamada A, Isawa H, Sawabe K, Omori S, Murata K, Yukawa M (2010) Molecular identification of blood source animals from black flies (Diptera: Simuliidae) collected in the alpine regions of Japan. *Parasitol Res* 106:543–547

- Ito A, Gotanda T (2004) Field efficacy of recombinant R7 vaccine against chicken Leucocytozoonosis. *J Vet Med Sci* 66:483–487
- Ito A, Gotanda T, Kobayashi S, Kume K, Sugimoto C, Matsumura T (2005) Increase of antibody titer against *Leucocytozoon caulleryi* by oral administration of recombinant R7 antigen. *J Vet Med Sci* 67: 211–213
- Janouskovec J, Horák A, Oborník M, Lukes J, Keeling PJ (2010) A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proc Natl Acad Sci U S A* 107:10949–10954
- Köhler S, Delwiche CF, Denny PW, Tilney LG, Webster P, Wilson RJM, Palmer JD, Roos DS (1997) A plastid of probable green algal origin in Apicomplexan parasites. *Science* 275:1485–1489
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948
- Lau AO, McElwain TF, Brayton KA, Knowles DP, Roalson EH (2009) *Babesia bovis*: a comprehensive phylogenetic analysis of plastid-encoded genes supports green algal origin of apicoplasts. *Exp Parasitol* 123:236–243
- McFadden GI (2011) The apicoplast. *Protoplasma* 248:641–650
- Morii T (1992) A review of *Leucocytozoon caulleryi* infection in chickens. *J Protozool Res* 2:128–133
- Morii T, Shihara T, Lee YC, Manuel MF, Nakamura K, Iijima T, Hoji K (1981) Seroimmunological and parasitological surveys of *Leucocytozoon caulleryi* infection in chickens in several Asian countries. *Int J Parasitol* 11:187–190
- Morii T, Matsui T, Fujino T, Kobayashi F, Tsuji T (1996) Immunogenicity of *Leucocytozoon caulleryi* sporozoites and their reactivity with specific immune sera. *Parasitol Res* 82:454–458
- Oborník M, Jirků M, Slapeta JR, Modrý D, Koudela B, Lukes J (2002) Notes on coccidian phylogeny, based on the apicoplast small subunit ribosomal DNA. *Parasitol Res* 88:360–363
- Omori S, Sato Y, Hirakawa S, Isobe T, Yukawa M, Murata K (2008) Two extra chromosomal genomes of *Leucocytozoon caulleryi*; complete nucleotide sequences of the mitochondrial genome and existence of the apicoplast genome. *Parasitol Res* 103:953–957
- Ortego J, Cordero PJ (2009) PCR-based detection and genotyping of haematozoa (Protozoa) parasitizing eagle owls, *Bubo bubo*. *Parasitol Res* 104(2):467–470
- Ralph SA, van Dooren GG, Waller RF, Crawford MJ, Fraunholz MJ, Foth BJ, Tonkin CJ, Roos DS, McFadden GI (2004) Tropical infectious diseases: metabolic maps and functions of the *Plasmodium falciparum* apicoplast. *Nat Rev Microbiol* 2:203–216
- Sato S, Sesay AK, Holder AA (2013) The unique structure of the apicoplast genome of the rodent malaria parasite *Plasmodium chabaudi chabaudi*. *PLoS One* 8:e61778. doi:10.1371/journal.pone.0061778
- Valkiunas G (2005) Avian malaria parasites and other Haemosporidia. CRC, Boca Raton
- Williamson DH, Denny PW, Moore PW, Sato S, McCready S, Wilson RJM (2001) The in vivo conformation of the plastid DNA of *Toxoplasma gondii*: implications for replication. *J Mol Biol* 306: 159–168
- Williamson DH, Preiser PR, Moore PW, McCready S, Strath M, Wilson RJM (2002) The plastid DNA of the malaria parasite *Plasmodium falciparum* is replicated by two mechanisms. *Mol Microbiol* 45: 533–542
- Wilson RJM, Denny PW, Preiser PR, Rangachari K, Roberts K, Roy A, Whyte A, Strath M, Moore DJ, Moore PW, Williamson DH (1996) Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. *J Mol Biol* 261:155–172
- Yeh E, DeRisi JL (2011) Chemical rescue of malaria parasites lacking an apicoplast defines organelle function in blood-stage *Plasmodium falciparum*. *PLoS Biol* 9:e1001138. doi:10.1371/journal.pbio.1001138
- Yu CY, Wang JS (2001) Role of chicken serum in inhibiting *Leucocytozoon caulleryi* development in *Culicoides arakawae* infected by membrane-feeding of infective blood meals. *Parasitol Res* 87:698–701