Parasitology

cambridge.org/par

De novo synthesis of thiamine (vitamin B_1) is the ancestral state in *Plasmodium* parasites – evidence from avian haemosporidians

Olof Hellgren, Staffan Bensch and Elin Videvall

MEMEG, Department of Biology, Lund University, Sölvegatan 37, 223 62 Lund, Sweden

Research Article

Cite this article: Hellgren O, Bensch S, Videvall E (2018). *De novo* synthesis of thiamine (vitamin B_1) is the ancestral state in *Plasmodium* parasites – evidence from avian haemosporidians. *Parasitology* **145**, 1084–1089. https://doi.org/10.1017/S0031182017002219

Received: 29 June 2017
Revised: 8 November 2017
Accepted: 9 November 2017
First published online: 12 December 2017

Key words:
Apicomplexa; auxotrophy; avian malaria; gene loss; plasmodium; thiamine

Author for correspondence:
Olof Hellgren, Email: Olof.Hellgren@biol.lu.se

Abstract

Parasites often have reduced genomes as their own genes become redundant when utilizing their host as a source of metabolites, thus losing their own *de novo* production of metabolites. Primate malaria parasites can synthesize vitamin B_1 (thiamine) *de novo* but rodent malaria and other genome-sequenced apicomplexans cannot, as the three essential genes responsible for this pathway are absent in their genomes. The unique presence of functional thiamine synthesis genes in primate malaria parasites and their sequence similarities to bacterial orthologues, have led to speculations that this pathway was horizontally acquired from bacteria. Here we show that the genes essential for the *de novo* synthesis of thiamine are found also in avian *Plasmodium* species. Importantly, they are also present in species phylogenetically basal to all mammalian and avian *Plasmodium* parasites, i.e. *Haemoproteus*. Furthermore, we found that these genes are expressed during the blood stage of the avian malaria infection, indicating that this metabolic pathway is actively transcribed. We conclude that the ability to synthesize thiamine is widespread among haemosporidians, with a recent loss in the rodent malaria species.

Introduction

When organisms evolve a parasitic life strategy, this often results in gene losses and size reduction of the genomes (Tamas *et al.* 2001; Sakharkar *et al.* 2007; Lee and Marx, 2012). This is mainly because parasites utilize metabolites and proteins from their hosts, making some metabolic machineries redundant (Kemen *et al.* 2011). However, different host–parasite combinations and life history traits will cause the loss of different genes and pathways, and in some cases parasites might even acquire or evolve new metabolic pathways to complement those of its host (Pombert *et al.* 2012).

In most organisms, thiamine (vitamin B_1) acts as an essential co-factor for several enzymes involved in carbohydrate and amino acid metabolism. Fungi, bacteria and plants can synthesize thiamine *de novo*, whereas most other organisms must salvage it from external sources.

De novo synthesis of thiamine requires enzymes that produce two essential moieties [4-amino-2-methyl-5-diphosphomethylpyrimidine and 4-methyl-5-(2-phosphoethyl)-thiazole), as well as an enzyme that combines these moieties into thiamine monophosphate (Helliwell *et al.* 2013, Fig. 1). In the primate malaria parasites, *Plasmodium falciparum*, *Plasmodium vivax* and *Plasmodium knowlesi*, the genes coding for these three key enzymes (hydroxyethylthiazole kinase ThiM, hydroxymethylpyrimidine kinase ThiD and thiamine-phosphate diphosphorylase ThiE) have been found (Frech and Chen, 2011). Further, in *P. falciparum*, the proteins coded by the genes ThiM, ThiD and ThiE have been shown experimentally to exhibit the expected enzymatic functions in the thiamine pathway (Wrenger *et al.* 2006). However, these enzymes have been shown to be absent in the rodent malaria parasites (*Plasmodium berghei*, *Plasmodium chabaudi* and *Plasmodium yoelii*) (Frech and Chen, 2011). Likewise, in the sequenced genomes of apicomplexan parasites more distantly related to malaria parasites (i.e. *Theileria*, *Babesia*, *Toxoplasma* spp., *Neospora* spp., *Eimeria* spp. and *Cryptosporidia* spp.), none of the required enzymes for this *de novo* synthesis have been found (Müller and Kappes, 2007; Shanmugasundram *et al.* 2013). Although thiamine is essential for survival (Helliwell *et al.* 2013), loss of this pathway leads to thiamine auxotrophy (i.e. vitamin dependence from an external source) and has occurred repeatedly in the course of evolution in eukaryotes, suggesting a trade-off in the cost of synthesizing the vitamin against the probability of salvaging it externally. The absence of the pathway in all apicomplexan parasites except *Plasmodium* parasites infecting primates has led to the speculation that the essential genes for the pathway in primate *Plasmodium* were retrieved by horizontal gene transfer from a bacterium (Frech and Chen, 2011).

Malaria parasites are not exclusive to rodents and primates, however, as many related *Plasmodium* species exist in lizards, bats and birds (Valkiunas, 2005; Martinsen *et al.* 2008; Schaer *et al.* 2013). The presence of the thiamine pathway in primate malaria species and its absence in rodent malaria parasites raises questions of its origin and possible evolutionary losses. Is the functional *de novo* thiamine synthesis pathway exclusive to *Plasmodium* parasites infecting primates? Or is thiamine synthesis a common trait in other *Plasmodium* parasites but has for unknown reasons been lost in rodent malaria parasites? In order to gain insights into

CAMBRIDGE UNIVERSITY PRESS

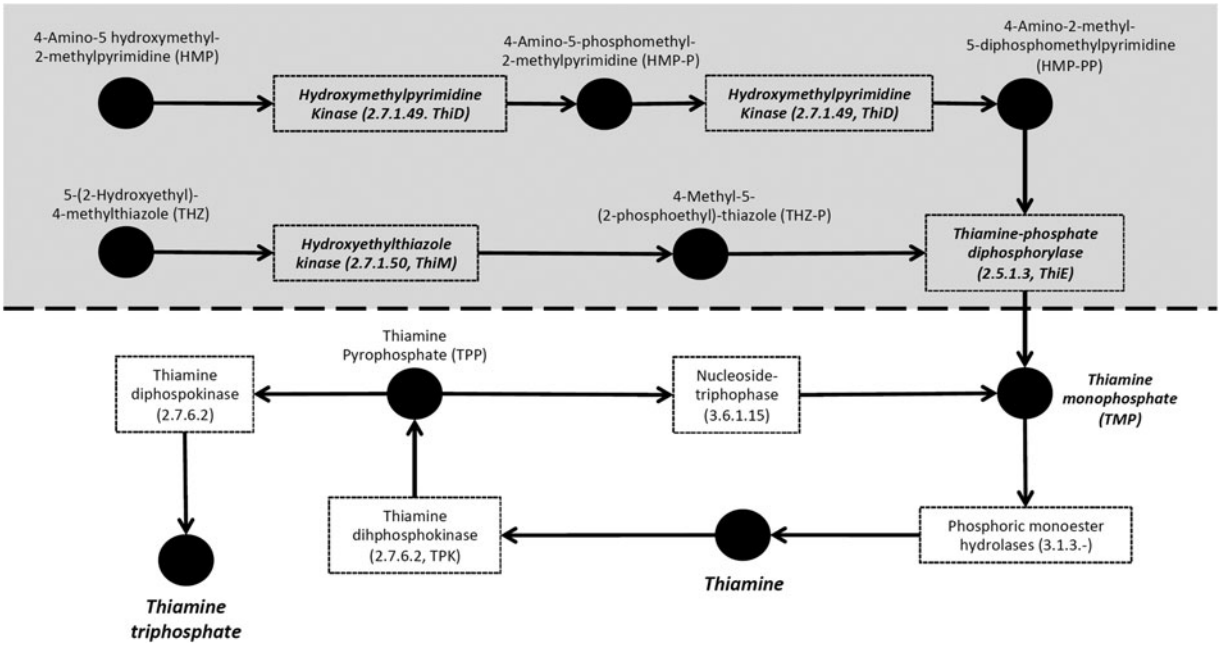

Fig. 1. Thiamine metabolism of *haemosporidian* parasites. The shaded area represents the pathway involved in *de novo* synthesis of thiamine, i.e. the enzymes needed to produce thiamine monophosphate. Boxes represent enzymes, and names outside boxes represent the products. Enzymes in bold and italic represent enzymes found in the genome of primate malaria parasites and the avian haemosporidian parasites *Haemoproteus tartakovskyi*, *Plasmodium relictum* and *Plasmodium ashfordi*.

the evolutionary origin of the thiamine pathway in *Plasmodium* parasites, we searched for orthologous thiamine genes in two transcriptomes of the avian malaria parasites *Plasmodium ashfordi* and *Plasmodium relictum*, and in the genomes of the avian parasites *Plasmodium gallinaceum* and *Haemoproteus tartakovskyi*. The existence of this pathway in the avian *Plasmodium* and *Haemoproteus* parasites could indicate that *de novo* thiamine synthesis is the ancestral state of *Plasmodium* parasites and that its absence in the rodent malaria species is a secondary loss. To get a more detailed picture of evolutionary gains and losses of the thiamine pathway in the malaria parasite clade, we extended the search for thiamine genes to several genomes of recently sequenced species of primate and rodent *Plasmodium*.

Method

Sequences of the single-exon genes that encode the three key enzymes necessary for *de novo* synthesis of thiamine in *P. falciparum*, hydroxyethylthiazole kinase (PFL1920c, EC 2.7.1.50), hydroxymethylpyrimidine kinase, (PFE1030c, EC 2.7.1.49) and thiamine-phosphate diphosphorylase (PFF0680c, EC 2.5.1.3), were downloaded from GenBank. They were then used in searches of orthologous genes in the two avian *Plasmodium* transcriptomes, *P. ashfordi* (Videvall *et al.* 2017), *P. relictum* and in the genome of *H. tartakovskyi* (Bensch *et al.* 2016). Because the *P. relictum* transcriptome is unpublished at present date, retrieved contigs were verified against the genome of *P. relictum* (Boehme *et al.* 2016, www.plasmodb.org) in order to validate that the contigs had assembled correctly. Searches were performed using local BLAST with the software Geneious ver. 6.1. Parameters for the local BLAST searches were set to: scoring (match mismatch): 2–3, seed length: 18, word size: 11, gap costs (open extend): 5 2. The contig with the lowest *E*-value was kept and used in subsequent analyses. The obtained sequences were translated and used in a BLAST search against the NCBI non-redundant protein database (BLASTP: BLOSOM62, word size 3, exp. threshold 10, gap cost 11 1. BLASTN: exp. threshold 10, word size 11, match mismatch scores 2,−3, gap costs 5 2) to ensure that they were true orthologues.

The genes from *P. falciparum* were further used to search the published mammalian malaria genomes of *Plasmodium malariae*, *Plasmodium coatneyi*, *Plasmodium inui*, *Plasmodium fragile*, *Plasmodium gaboni*, *Plasmodium reichenowi and Plasmodium vinckei*, retrieved from Plasmodb.org, release 34 (Aurrecoechea *et al.* 2009). Obtained genes from *P. relictum* were used to search the genome of *P. gallinaceum* in GeneDB.org (Logan-Klumpler *et al.* 2013).

Protein alignments were conducted using the MUSCLE-alignment algorithm, with 100 iterations and the default settings, as implemented in Geneious ver. 6.1.6.

Gene phylogenies of the obtained protein sequences for each of the genes were constructed using a maximum-likelihood method and each tree was resampled using 500 bootstrap iterations as implemented in MEGA7 (Kumar *et al.* 2016). Rates of molecular change were set to a JTT + G + F (with five discrete γ distributions) model for the genes ThiD and ThiE and a LG + G (with five discrete γ distributions) model for the gene ThiM. Models were selected based on the lowest Bayesian information criterion scores after running the model test as implemented in MEGA7 (Kumar *et al.* 2016).

Results

The orthologous genes encoding the three key enzymes involved in *de novo* synthesis of thiamine (hydroxyethylthiazole kinase, hydroxymethylpyrimidine kinase and thiamine-phosphate diphosphorylase) were found in all three newly sequenced unannotated avian haemosporidians *P. relictum*, *P. ashfordi* and *H. tartakovskyi* (GenBank nr: KP784838, KP784836, KP78483, KP784835, KP784833, KP784837, KP784841, KP784839, KP784841), and in the genome of the annotated avian parasite *P. gallinaceum* (GeneDB GeneID: PGAL8A_00461300, PGAL8A_00067800, PGAL8A_00142600). Local Blast of the three avian parasites *P. relictum*, *P. ashfordi* and *H. tartakovskyi* yielded highly significant matches (*E*-values $<9{\cdot}3 \times 10^{-31}$) (Table 1). The second best hits had considerably higher *E*-values (ranges between 1×10^{-7} and 1×10^{-2}, Appendix 1). Gene annotations of the retrieved sequences were confirmed with BLAST searches against

https://doi.org/10.1017/S0031182017002219 Published online by Cambridge University Press

Table 1. Local Blast results of the best hits of the genes ThiD, ThiE and ThiM from *Plasmodium falciparum* against *Plasmodium ashfordi* (GRW2), *Plasmodium relictum* (SGS1) and *Haemoproteus tartakovskyi* (SISKIN1)

Protein	*P. falciparum* Gene ID	*P. ashfordi* *E*-value	*P. relictum* *E*-value	*H. tartakovskyi* *E*-value
Hydroxymethylpyrimidine kinase (ThiD)	PFE1030c	$2{\cdot}9 \times 10^{-47}$	$7{\cdot}6 \times 10^{-55}$	$9{\cdot}3 \times 10^{-31}$
Thiamine-phosphate diphosphorylase (ThiE)	PFF0680c	$3{\cdot}4 \times 10^{-49}$	$4{\cdot}6 \times 10^{-92}$	$8{\cdot}5 \times 10^{-28}$
Hydroxyethylthiazole kinase (ThiM)	PFL1920c	$1{\cdot}7 \times 10^{-151}$	$<1 \times 10^{-179}$	3×10^{-100}

GenBank, which yielded highly significant *E*-values both for the nucleotide sequences (between 1×10^{-32} and $<1 \times 10^{-100}$) and for the translated sequences (1×10^{-46} to 1×10^{-164}), all matching the expected annotated genes (Appendices 2 and 3). These results strongly indicate that the orthologous thiamine genes have been identified correctly in *P. ashfordi*, *P. relictum* and *H. tartakovskyi*. Within the genomes of *P. malariae*, *P. coatneyi*, *P. inui*, *P. fragile*, *P. gaboni* and *P. reichenowi*, all three genes were located and the orthology was confirmed with conserved protein alignments (Supplementary Fig. S1a–c) and phylogenetic clustering with the previously annotated genes (Fig. 2). However, none of the thiamine genes were found in the genome of the newly sequenced rodent malaria parasite *P. vinckei*, which together with *P. yoelii*, *P. berghei and P. chabaudi* (Frech and Chen, 2011), makes it a total of four rodent malaria species discovered without these genes.

All protein sequence alignments showed large regions with highly conserved blocks of amino acids (Supplementary Fig. S1a–c) as well as an overall high level of sequence similarity, further strengthening the case that we had received the orthologues of the genes coding for the enzymes in the thiamine biosynthesis pathway.

In no cases did the retrieved orthologous genes contain stop codons within the proposed exons. The transcriptomes of *P. relictum* and *P. ashfordi* were derived from the erythrocytic phase of infection (for methods see Videvall *et al.* 2015), and the presence of expressed thiamine genes during this stage demonstrates that the genes not only are present in these genomes, but are indeed activated and transcribed during this part of the infection cycle. The full length of the genes was obtained for *P. relictum* and *H. tartakovskyi*, whereas the ThiD and THiE transcripts from *P. ashfordi* were slightly shorter. The shorter transcripts are most likely an artefact due to insufficient sequence coverage of the *P. ashfordi* transcriptome, which have resulted in a slightly lower mean and median length of the transcripts in comparison to *P. falciparum* (Videvall *et al.* 2017).

Discussion

In apicomplexan parasites, the metabolic pathway for *de novo* synthesis of thiamine is not exclusive to malaria parasites infecting primates (Table 2, Fig. 3). This conclusion is based on our finding of the key enzymes in the thiamine pathway being expressed in two

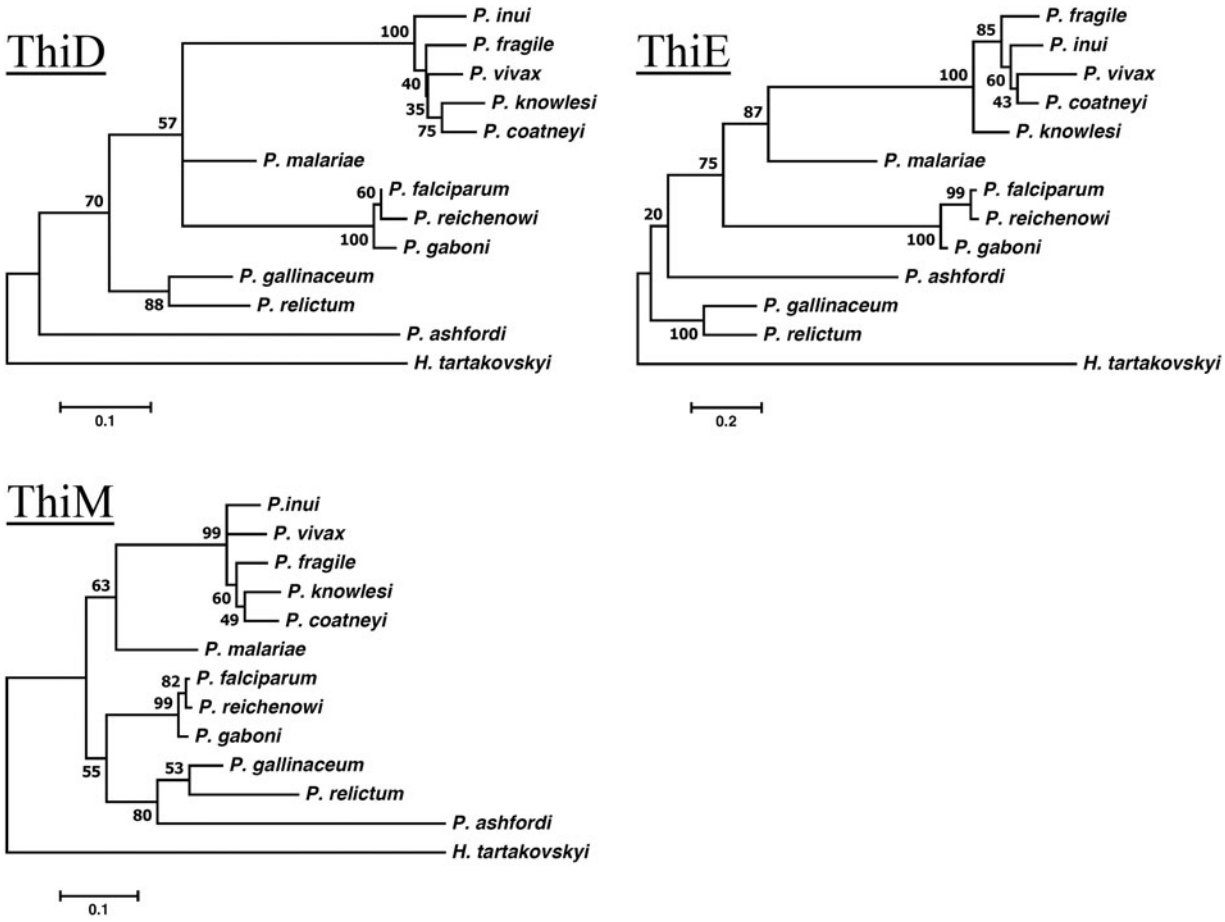

Fig. 2. Maximum-likelihood phylogenies of the ThiD, ThiE and ThiM thiamine genes. Numbers on branches represent bootstrap values based on 500 iterations.

https://doi.org/10.1017/S0031182017002219 Published online by Cambridge University Press

Table 2. Presence (+) or absence (–) of the genes ThiD, ThiM and ThiE in *Plasmodium* and *Haemoproteus* species

Species	ThiD	ThiM	ThiE	Hosts
P. falciparum	+	+	+	Primates
P. vivax	+	+	+	Primates
P. malariae	+	+	+	Primates
P. reichenowi	+	+	+	Primates
P. gaboni	+	+	+	Primates
P. knowlesi	+	+	+	Primates
P. inui	+	+	+	Primates
P. coatneyi	+	+	+	Primates
P. fragile	+	+	+	Primates
P. ashfordi	+	+	+	Birds
P. relictum	+	+	+	Birds
P. gallinaceum	+	+	+	Birds
P. berghei	–	–	–	Rodents
P. yoelii	–	–[a]	–	Rodents
P. chabaudi	–	–	–	Rodents
P. vinckei	–	–	–	Rodents
H. tartakovskyi	+	+	+	Birds

Species in bold represent findings presented in this study, whereas non-bold species represent previously reported findings from Frech and Chen (2011).
[a]Represents a degenerated gene fragment of ThiM found in *P. yoelii*.

avian *Plasmodium* species and their gene orthologues present in the genomes of a third avian *Plasmodium* species and a parasite in the sister genus *Haemoproteus*.

There are several competing hypotheses regarding the phylogeny of *Plasmodium* species, and in particular the evolutionary relationship between *P. falciparum* and avian *Plasmodium* species (Perkins, 2014). However, there is a strong consensus that the rodent malaria parasites form a single monophyletic clade (Martinsen *et al.* 2008; Pick *et al.* 2011; Bensch *et al.* 2016; Borner *et al.* 2016). Further, recent phylogenetic analyses based on genome sequencing of *H. tartakovskyi* show that *Haemoproteus* is a sister taxon to a monophyletic clade of all *Plasmodium* species (Bensch *et al.* 2016). Thus, our findings indicate that the thiamine pathway is ancestral to the whole group of *Plasmodium* as they are found in the sister genus *Haemoproteus* as well. In the genomes of rodent malaria parasites, a remnant of these genes can be seen in *P. yoelii*, which has one gene relict of the hydroxyethylthiazole kinase (ThiM) (Frech and Chen, 2011). The reason why rodent *Plasmodium* species have lost the genes involved in the thiamine synthesis pathway remains unknown. The four rodent malaria parasites that have been sequenced are all derived from isolates kept for a number of generations in laboratory mice fed on grain which is rich in thiamine. Loss of the thiamine genes due to such artificial selection is possible, however unlikely, as it would require parallel independent losses in all four species. A more parsimonious explanation is that the loss happened prior to the radiation of the rodent malaria parasites.

Haemoproteus and avian *Plasmodium* are extremely species-rich genera of pathogens (Bensch *et al.* 2009), with numerous and mainly unstudied host–pathogen combinations. These parasites might therefore prove to be suitable systems for investigating repeated gene losses in order to find common evolutionary denominators for when this essential pathway is lost. It would be of particular interest to investigate the presence of the thiamine pathway in avian malaria species that are specialists on granivore and insectivore hosts, to elucidate whether gene losses can be associated with the hosts' food preferences.

Frech and Chen (2011) speculated that the three key enzymes for thiamine synthesis in primate malaria parasites originated *via*

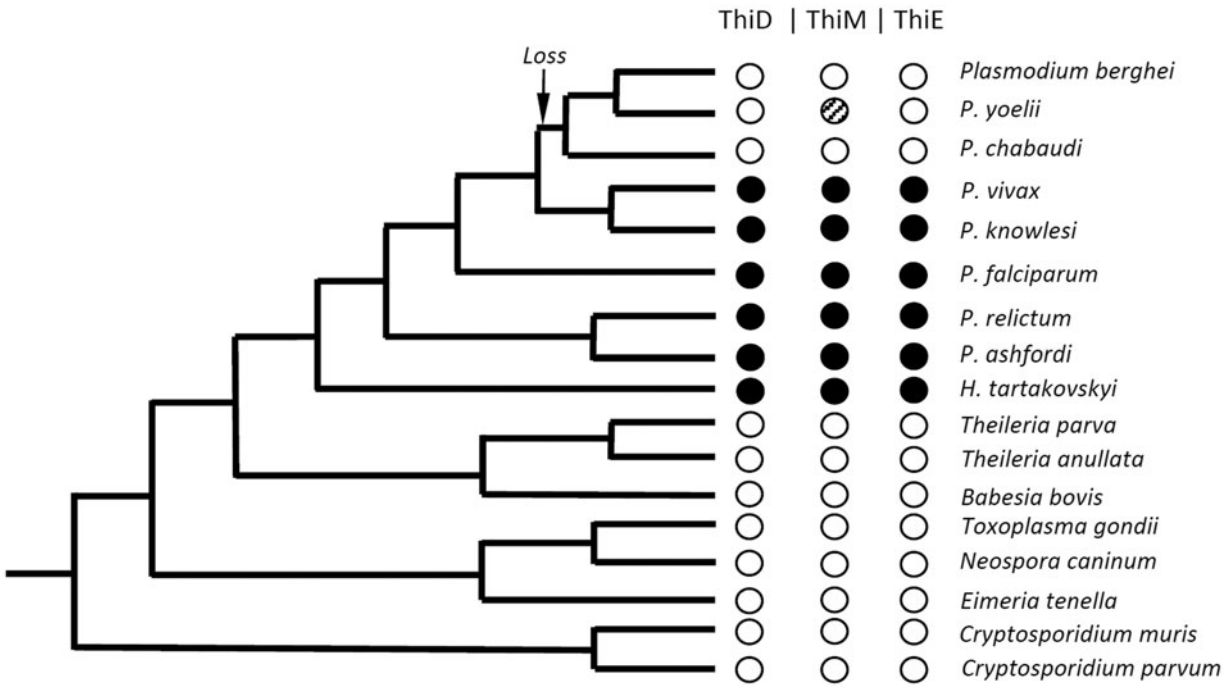

Fig. 3. Cladogram of selected apicomplexan parasites. Black dots represent the presence, and white dots represent the absence of the thiamine enzymes ThiD, ThiM and ThiE. The shaded dot represents a case where a non-functional gene remnant is present in the genome. The loss arrow represents the case where the genes have been lost before the species splits seen today. The cladogram is redrawn based on several published phylogenies of apicomplexan parasites (Martinsen *et al.* 2008; Kissinger and DeBarry, 2011; Weatherby and Carter, 2013; Bensch *et al.* 2016; Borner *et al.* 2016). Note that the branch lengths are not to scale but are an illustration of the phylogenetic relationship between the parasites. For illustrative purposes, only the parasites with known phylogenetic placements are used in the cladogram.

https://doi.org/10.1017/S0031182017002219 Published online by Cambridge University Press

horizontal gene transfer from a bacterium. They found that these genes had both sequence similarities with *Clostridium* spp. and their genomic location next to each other on the same strand in the *Clostridium* genome enabled the formation of a potential operon (Frech and Chen, 2011). If these genes have been acquired through horizontal gene transfer, then this event must have happened prior to the split of the genera *Haemoproteus* and *Plasmodium* (Fig. 3), thus being ancestral to all *Plasmodium* species. On the other hand, we cannot exclude that these genes constitute the ancestral stage of all apicomplexans, followed by repeated losses throughout the phylogeny but being kept within the clade of the haemosporidian species (Fig. 3).

The presence of the *de novo* pathway of thiamine across primate *Plasmodium* species have made it a potential target for antimalarial drugs research (Chan *et al.* 2013). As this metabolic pathway does not exist in the most important model organisms in malaria research, rodent *Plasmodium*, it has been difficult to explore the potential for using the thiamine pathway in drug development. In the early days of malaria research, birds were frequently used as model organisms and served in some of the most fundamental breakthroughs in malaria research and drug development (von Wasielewski, 1904–08; McGhee *et al.* 1977; Cox, 2010). The presence of the thiamine pathway in avian malaria species highlights the value of studying birds as a complement model system in the searches for antimalarial drug targets.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0031182017002219.

Authors' contributions. OH conceived the study, performed the analysis and wrote the first draft of the manuscript. EV performed the bioinformatic work on *P. ashfordi* and *P. relictum*. SB performed the bioinformatic work on *H. tartakovskyi*. All authors participated in interpreting the data, contributed to the writing and approved of the final manuscript. Genbank accession numbers of the ThiD, ThiM and ThiE genes from *P. ashfordi*, *P. relictum* and *H. tartakovskyi* have been submitted by the authors and are to be public upon publication.

Funding. This research was founded by the Swedish research council (VR) and The Crafoord foundation.

Competing interests. The authors declare to have no financial, non-financial competing interests or competing interest from commercial organizations.

Availability of data and material. All data used in the study have been published in public databases and can be available from the authors on reasonable request.

References

Aurrecoechea C, Brestelli J, Brunk BP, Dommer J, Fischer S, Gajria B, Gao X, Gingle A, Grant G, Harb OS, Heiges M, Innamorato F, Iodice J, Kissinger JC, Kraemer E, Li W, Miller JA, Nayak V, Pennington C, Pinney DF, Roos DS, Ross C, Stoeckert CJ Jr, Treatman C and Wang H (2008) PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acid Research* **37**, D539–D543.

Bensch S, Hellgren O and Pérez-Tris J (2009) Malavi: a public database of malaria parasites and related haemosporidians in avian hosts based on mitochondrial cytochrome b lineages. *Molecular Ecology Resources* **9**, 1353–1358.

Bensch S, Canbäck B, DeBarry JD, Johansson T, Hellgren O, Kissinger JC, Palinauskas V, Videvall E and Valkiūnas G (2016) The genome of *Haemoproteus tartakovskyi* and its relationship to human malaria parasites. *Genome Biology and Evolution* **8**, 1361–1373.

Boehme U, Otto TD, Cotton J, Steinbiss S, Sanders M, Oyola SM, Nicot A, Gandon S, Patra KP, Herd C, Bushell E, Modrzynska KK, Billker O, Vinetz JM, Rivero A, Newbold CI and Berriman M (2016) Complete avian malaria parasite genomes reveal host-specific parasite evolution in birds and mammals. *bioRxiv*. https://doi.org/101.101/086504.

Borner J, Pick C, Thiede J and Kolawole OM (2016) Phylogeny of haemosporidian blood parasites revealed by a multi-gene approach. *Molecular Phylogenetics and Evolution* **94**, 221–231.

Chan XWA, Wrenger C, Stahl K, Bergmann B, Winterberg M, Müller IB and Saliba KJ (2013) Chemical and genetic validation of thiamine utilization as an antimalarial drug target. *Nature Communications* **4**, 2060.

Cox FE (2010) History of the discovery of the malaria parasites and their vectors. *Parasites & Vectors* **3**, 5.

Frech C and Chen N (2011) Genome comparison of human and non-human malaria parasites reveals species subset-specific genes potentially linked to human disease. *PLoS Computational Biology* **7**, e1002320.

Helliwell KE, Wheeler GL and Smith AG (2013) Widespread decay of vitamin-related pathways: coincidence or consequence? *Trends in Genetics* **29**, 469–478.

Kemen E, Gardiner A, Schultz-Larsen T, Kemen AC, Balmuth AL, Robert-Seilaniantz A, Bailey K, Holub E, Studholme DJ, MacLean D and Jones JDG (2011) Gene gain and loss during evolution of obligate parasitism in the white rust pathogen of Arabidopsis thaliana. *PLoS Biology* **9**, e1001094.

Kissinger JC and DeBarry J (2011) Genome cartography: charting the apicomplexan genome. *Trends in Parasitology* **27**, 345–354.

Kumar S, Stecher G and Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**, 1870–1874.

Lee M.-C and Marx CJ (2012) Repeated, selection-driven genome reduction of accessory genes in experimental populations. *PLoS Genetics* **8**, e1002651.

Logan-Klumpler FJ, De SIlva N, Boehme U, Rogers MB, Vilarde G, McQuillan JA, Carver T, Aslett M, Olsen C, Phan SSI, Farris C, Mitra S, Ramasamy G, Wang H, Tivey A, Jackson A, Houston R, Parkhill J, Holden M, Harb OS, Brunk BP, Myler PJ, Roos D, Carrington M, Smith DF, Hertz-Fowler C and Berriman M (2012) GeneDB-an annotation database for pathogens. *Nucleic Acid Research* **40**, D98–D108. https://doi.org/10.1093/nar/gkr1032.

Martinsen ES, Perkins SL and Schall JJ (2008) A three-genome phylogeny of malaria parasites (*Plasmodium* and closely related genera): evolution of life-history traits and host switches. *Molecular Phylogenetics and Evolution* **47**, 261–273.

McGhee RB, Singh SD and Weathersby AB (1977) *Plasmodium gallinaceum*: vaccination in chickens. *Experimental Parasitology* **43**, 231–238.

Müller S and Kappes B (2007) Vitamin and cofactor biosynthesis pathways in *Plasmodium* and other apicomplexan parasites. *Trends in Parasitology* **23**, 112–121.

Perkins SL (2014) Malaria's many mates: past, present, and future of the systematics of the order *Haemosporida*. *The Journal of Parasitology* **100**, 11–25.

Pick C, Ebersberger I, Spielmann T, Bruchhaus I and Burmester T (2011) Phylogenomic analyses of malaria parasites and evolution of their exported proteins. *BMC Evolutionary Biology* **11**, 167.

Pombert J.-F, Selman M, Burki F, Bardell FT, Farinelli L, Solter LF, Whitman DW, Weiss LM, Corradi N and Keeling PJ (2012) Gain and loss of multiple functionally related, horizontally transferred genes in the reduced genomes of two microsporidian parasites. *Proceedings of the National Academy of Sciences of the USA* **109**, 12638–12643.

Sakharkar KR, Sakharkar MK and Chow VTK (2007) Gene loss and gene conservation in obligatory intracellular parasites of human. *Current Bioinformatics* **2**, 214–221.

Schaer J, Perkins SL, Decher J, Leendertz FH, Fahr J, Weber N and Matuschewski K (2013) High diversity of West African bat malaria parasites and a tight link with rodent Plasmodium taxa. *Proceedings of the National Academy of Sciences* **110**, 17415–17419.

Shanmugasundram A, Gonzalez-Galarza FF, Wastling JM, Vasieva O and Jones AR (2013) Library of apicomplexan metabolic pathways: a manually curated database for metabolic pathways of apicomplexan parasites. *Nucleic Acids Research* **41**, D706–D713.

Tamas I, Klasson LM, Sandstrom JP and Andersson S (2001) Mutualists and parasites: how to paint yourself into a (metabolic) corner. *Febs Letters* **498**, 135–139.

Valkiunas G (ed.) (2005) *Avian Malaria Parasites and Other Haemosporidia*. Boca Raton, Florida: CRC.

Videvall E, Cornwallis CK, Palinauskas V, Valkiūnas G and Hellgren O (2015) The avian transcriptome response to malaria infection. *Molecular Biology and Evolution* **32**, 1255–1267.

Videvall E, Cornwallis CK, Ahren D, Palinauskas V, Valkiūnas G and Hellgren O (2017) The transcriptome of the avian malaria parasite *Plasmodium ashfordi* displays host-specific gene expression. *Molecular Ecology* **26**, 2939–2958.

von Wasielewski TKWN (1904–08) *Studien und Mikrophotogramme zur Kenntnis der pathogenen Protozoen*. Leipzig: Barth.

Weatherby K and Carter D (2013) *Chromera velia*: the missing link in the evolution of parasitism. *Advances in Applied Microbiology* **85**, 119–144.

Wrenger C, Eschbach M-L, Müller IB, Laun NP, Begley TP and Walter RD (2006) Vitamin B1 de novo synthesis in the human malaria parasite *Plasmodium falciparum* depends on external provision of 4-amino-5-hydroxymethyl-2-methylpyrimidine. *Biological Chemistry* **387**, 41–51.

Appendix 1

Table A1. Local Blast results of the second best hits of the genes ThiD, ThiE and ThiM from *Plasmodium falciparum* against genes in the transcriptome and genome of *Plasmodium ashfordi* (GRW2), *Plasmodium relictum* (SGS1) and *Haemoproteus tartakovskyi* (*SISKIN1*)

Protein	*P. falciparum* Gene ID	*P. ashfordi* *E*-value	*P. relictum* *E*-value	*H. tartakovskyi* *E*-value
Hydroxymethylpyrimidine kinase (ThiD)	PFE1030c	1.8×10^{-5}	8.1×10^{-4}	1.6×10^{-2}
Thiamine-phosphate diphosphorylase (ThiE)	PFF0680c	4.8×10^{-3}	7.86×10^{-7}	2.3×10^{-3}
Hydroxyethylthiazole kinase (ThiM)	PFL1920c	2.7×10^{-3}	7.9×10^{-4}	1.0×10^{-4}

Appendix 2

Table A2. nBLAST (exp threshold 10, word size 11, match mismatch scores 2, −3, gap costs 5 2) and pBLAST (BLOSOM62, word size 3, exp. threshold 10, gap cost 11 1) results of putative ThiD, ThiE and ThiM genes obtained from *Plasmodium ashfordi* (GRW2), *Plasmodium relictum* (SGS1) and *Haemoproteus tartakovskyi* (*SISKIN1*) against GenBank

Protein	*P. ashfordi* *E*-value (n/p)	*P. relictum* *E*-value (n/p)	*H. tartakovskyi* *E*-value (n/p)
Hydroxymethylpyrimidine kinase (ThiD)	$3 \times 10^{-101}/3 \times 10^{-113}$	$2 \times 10^{-179}/1 \times 10^{-156}$	$1 \times 10^{-45}/4 \times 10^{-132}$
Thiamine-phosphate diphosphorylase (ThiE)	$9 \times 10^{-71}/1 \times 10^{-80}$	$8 \times 10^{-137}/7 \times 10^{-144}$	$2 \times 10^{-32}/4 \times 10^{-45}$
Hydroxyethylthiazole kinase (ThiM)	$8 \times 10^{-147}/1 \times 10^{-142}$	$<1 \times 10^{-179}/9 \times 10^{-164}$	$3 \times 10^{-96}/1 \times 10^{-133}$

Appendix 3

BLASTn results of retrieved genes in the transcriptomes and genome of *Plasmodium relictum* (*SGS1*), *Plasmodium ashfordi* (GRW2) and *Haemoproteus tartakovskyi* (*Ht*).

GRW2_hydroxyethylthiazole kinase (KP784835)
Plasmodium falciparum 3D7 hydroxyethylthiazole kinase, putative (PFL1920c) mRNA, complete cds, XM_001350754.1, $e = 8 \times 10^{-147}$

GRW2_hydroxylmethylpyrimidine kinase KP784833
Plasmodium falciparum 3D7 phosphomethylpyrimidine kinase, putative (PFE1030c) mRNA, complete cds, XM_001351727.1, $e = 3 \times 10^{-101}$

GRW2_Thiamine-phosphate dihydrophosphorylase KP784837
Plasmodium falciparum 3D7 thiamine-phosphate pyrophosphorylase, putative (PFF0680c) mRNA, complete cds, XM_961034.2, $e = 9 \times 10^{-71}$

Ht_hydroxyethylthiazole kinase KP784841
Plasmodium falciparum 3D7 hydroxyethylthiazole kinase, putative (PFL1920c) mRNA, complete cds, XM_001350754.1, $e = 3 \times 10^{-96}$

Ht_hydroxylmethylpyrimidine kinase KP784839
Plasmodium falciparum 3D7 phosphomethylpyrimidine kinase, putative (PFE1030c) mRNA, complete cds, XM_001351727.1, $e = 1 \times 10^{-45}$

Ht_Thiamine-phosphate dihydrophosphorylase KP784841
Plasmodium falciparum 3D7 thiamine-phosphate pyrophosphorylase, putative (PFF0680c) mRNA, complete cds, XM_961034.2, $e = 2 \times 10^{-32}$

SGS1_hydroxyethylthiazole kinase KP784838
Plasmodium falciparum 3D7 hydroxyethylthiazole kinase, putative (PFL1920c) mRNA, complete cds, XM_001350754.1, $e = 0.0$

SGS1_hydroxylmethylpyrimidine kinase KP784836
Plasmodium falciparum 3D7 phosphomethylpyrimidine kinase, putative (PFE1030c) mRNA, complete cds, XM_001351727.1 $e = 2 \times 10^{-179}$

SGS1_Thiamine-phosphate dihydrophosphorylase KP78483
Plasmodium falciparum 3D7 thiamine-phosphate pyrophosphorylase, putative (PFF0680c) mRNA, complete cds, XM_961034.2, $e = 8 \times 10^{-137}$

https://doi.org/10.1017/S0031182017002219 Published online by Cambridge University Press