parasitologia

MDPI

Article

Duplication of a Type-P5B-ATPase in *Laverania* and Avian Malaria Parasites and Implications About the Evolution of *Plasmodium*

Mark F. Wiser

Department of Tropical Medicine and Infectious Disease, Tulane University Celia Scott Weatherhead School of Public Health & Tropical Medicine, New Orleans, LA 70112, USA; wiser@tulane.edu; Tel.: +1-504-988-2507

Abstract: Two related P-type ATPases, designated as ATPase1 and ATPase3, were identified in *Plasmodium falciparum*. These two ATPases exhibit very similar gene and protein structures and are most similar to P5B-ATPases. There are some differences in the predicted substrate-binding sites of ATPase1 and ATPase3 that suggest different functions for these two ATPases. Orthologues of ATPase3 were identified in all *Plasmodium* species, including the related *Hepatocystis* and *Haemoproteus*. ATPase3 orthologues could also be identified in all apicomplexan species, but no clear orthologues were identified outside of the Apicomplexa. In contrast, ATPase1 orthologues were only found in the *Laverania*, avian *Plasmodium* species, and *Haemoproteus*. ATPase1 likely arose from a duplication of the ATPase3 gene early in the evolution of malaria parasites. These results support a model in which early malaria parasites split into two clades. One clade consists of mammalian malaria parasites and *Hepatocystis* but excludes *P. falciparum* and related *Laverania*. The other clade includes *Haemoproteus*, avian *Plasmodium* species, and *Laverania*. This contrasts to recent models that suggest all mammalian malaria parasites form a monophyletic group, and all avian malaria parasites form a separate monophyletic group. ATPase1 may be a useful taxonomic/phylogenetic character for the phylogeny of Haemosporidia.

Keywords: P-type ATPase; *Plasmodium*; *Laverania*; *Haemoproteus*; Haemosporidia; Apicomplexa; malaria parasite; phylogeny; evolution

Academic Editor: Geoff Hide

Received: 10 December 2024
Revised: 10 January 2025
Accepted: 24 January 2025
Published: 27 January 2025

Citation: Wiser, M.F. Duplication of a Type-P5B-ATPase in *Laverania* and Avian Malaria Parasites and Implications About the Evolution of *Plasmodium*. *Parasitologia* **2025**, *5*, 6. https://doi.org/10.3390/parasitologia5010006

1. Introduction

P-type ATPases are part of a large and ubiquitous family of membrane proteins that move substances, especially cations or lipids, across biological membranes using the hydrolysis of ATP to provide the motive force [1]. The name P-type refers to transient phosphorylation of a specific aspartate residue during ATP hydrolysis. This switching between the phosphorylated and unphosphorylated states results in conformation changes that move substrates across the membranes. Thus, P-type ATPases are also known as E1–E2 ATPases in reference to the distinct molecular states that are involved in transport.

P-type ATPases are defined by distinct structural domains that include six core transmembrane helices (cTM), an actuator (A) domain, a phosphorylation (P) domain, and a nucleotide-binding (N) domain. Other domains that are not universally found in all P-type ATPases include a variable number of supporting transmembrane helices (sTM), N-terminal domains, or C-terminal domains. The helices of the cTM domain are numbered sequentially 1–6, and they form a channel that opens to the luminal or extracellular side

of the membrane and serves as the substrate-binding site. The fourth cTM helix is disrupted in the middle with a highly conserved proline, and this disruption, i.e., kink, in the helix forms part of the substrate-binding site. A conformation change during the phosphorylation–dephosphorylation cycles allows passage of the substrate across this kink to the cytoplasmic side of the membrane. The A-domain contains the intrinsic phosphatase activity and is composed of a sequence on the N-terminal side of cTM helix 1 (cTM1) and a sequence between cTM2 and cTM3. The P- and N-domains are formed by a large loop between cTM4 and cTM5. The P-domain consists of sequences closest to these two transmembrane helices, and the N-domain lies between the two parts of the P-domain. A defining feature of P-type ATPases is the phosphorylated aspartate (D) site, which has the highly conserved motif of DKTGT.

P-type ATPases are classified into six distinct types based on sequence homology, configuration of the sTM domains, presence or absence of N-terminal or C-terminal domains, and substrate specificity [1,2]. These types are numbered 1–6, and subtypes are designated with capital letters. The sequence homology is based on eight highly conserved regions, designated as cores A–H [2]. Core D is equivalent to cTM4. Outside of the core sequences and cTM helices, the different types of P-type ATPases exhibit distinct features. For example, P5-ATPases have four sTM helices on the C-terminal side of the cTM and additional membrane-associated helices on the N-terminal side of the cTM. These N-terminal membrane helices form a loop (nMHL) that consists of two or three short helices that do not traverse the lipid bilayer. In addition, between the nMHL and cTM1 is an N-terminal domain (NTD). P5-ATPases are divided into two subclasses designated as P5A and P5B [3].

Plasmodium falciparum has at least two P5-ATPases, which are designated as PfATPase1 [4] and PfATPase3 [5]. An orthologue of ATPase3 was also identified in *P. yoelii* [6]. Both of these ATPases are type-P5B [3]. Based on immunofluorescence, ATPase 3 is expressed predominantly in gametocytes in both *P. falciparum* [5] and *P. yoelii* [6]. However, ATPase3 mRNA expression is expressed in the asexual stages [5,7]. ATPase1 is also expressed in blood-stage asexual parasites [4]. Immunofluorescence studies have revealed a diffuse staining throughout the cytoplasm of the parasite for both ATPase1 and ATPase3 [4–6], which suggests an association with a cytoplasmic organelle.

Further characterization of these two *Plasmodium* ATPases reveal that ATPase3 is found in all apicomplexan species, whereas ATPase1 is only found in the subgenus *Laverania* and avian malaria parasites including *Haemoproteus*. *Laverania* consists of *P. falciparum* and several species of parasites that infect chimpanzees and gorillas [8]. The relationship of *Laverania* to mammalian or avian parasites has been long debated [9]. Characterization of the synteny, phylogeny, and structures of these two *Plasmodium* P5B-ATPases indicate that ATPase1 likely arose from a duplicative transposition of ATPase3 early in the evolution of the malaria parasite. These results support a shared common ancestor between *Laverania* and avian haemosporidian parasites.

2. Results

2.1. ATPase1

The first P-type ATPase to be cloned and sequenced from the human malaria parasite *P. falciparum* was designated as PfATPase1 (Genebank Acc. No. X65738) [4]. Complete orthologues of this gene were identified from seven species of the *Laverania* subgenus, the avian parasites *P. gallinaceum* and *P. relictum*, and *Haemoproteus tartakovskyi*. However, no orthologues of ATPase1 were found in other mammalian malaria parasites or *Hepatocystis*. There were inconsistencies in the proposed introns of the genes from *P. billcollinsi* and *P. praefalciparum,* and these sequences were not used in alignments. Alignment of the other

five *Laverania* sequences, the two avian parasite sequences, and the *H. tartakovskyi* sequence revealed common features of P-type ATPases, including the eight P-type ATPase sequence cores (A–H) and the six core transmembrane helices (Supplemental Figure S1). As previously noted [5], ATPase1 is most similar to the type-P5-ATPases due to the configuration of the supporting transmembrane helices and conserved amino acids. A membrane associated loop is found at the N-terminus and four supporting transmembrane helices are found at the C-terminus. In addition, ATPase1 has an extended phosphorylation sequence motif (FDKTGT[L/I]T), typical of the type-P5-ATPases [2], and ATPase1 has a PPxxP motif in the kink of cTM4, as is observed in many type-P5-ATPases [1]. In general, the conversed regions are most consistent with the type-P5B-ATPases, as previously noted [3].

Four rather large variable regions on the cytoplasmic face of the protein are also found within the ATPase1 sequence (Supplemental Figure S1). The greatest amount of variation is observed between the *Laverania* clade and the avian parasite clade including *Haemoproteus*. Within the clades, there is substantial sequence homology. As is often the case [10–12], these variable regions exhibit a high level of low-complexity sequence and tandem repeats. The first variable region is between the N-terminal membrane helical loop (nMHL) and the first core transmembrane helix (cTM1). It is characterized by tandem repeats in most of the *Laverania* sequences and an especially long insert of tandem repeats in *P. blacklocki*. Variable region 2 is found within the A-domain and occupies most of the sequence between sequence cores B and C. Variable region 3 is between core E and core F. There is a region of conserved sequence before core F that is interrupted by a small asparagine-rich insert of in the *Laverania* sequences. Variable region 4 is located between core F and core G, and a stretch of approximately 110 conserved amino acids is found within variable region 4.

2.2. ATPase3

An ATPase with similarities to ATPase1 was cloned and sequenced from *P. falciparum* [5] and *P. yoelii* [6] and designated as ATPase3. Rozmajzl et al. [5] identified the eight cores (A–H) found in P-type ATPases (Axelsen and Palmgren, 1998) and detected twelve predicted membrane-spanning helices [5]. ATPase3 orthologues are found in all *Plasmodium* species as well as *Hepatocystis*, and a partial sequence was identified in *Haemoproteus* (Appendix A). Alignment of the complete sequences reveals the typical features of P-type ATPases, including the eight core regions and the six core transmembrane helices (Supplemental Figure S2). In addition, the four supporting transmembrane helices at the C-terminus and the N-terminal membrane loop are also present. Originally, two supporting transmembrane helices at the N-terminal of PfATPase3 were predicted [5]. This difference in the previously predicted membrane helices is likely due to the more sophisticated programs that are now available and the increased knowledge about P-type ATPase structure. Furthermore, the predicted three-dimensional structures of ATPase3 are consistent, with two or three short helices at the N-terminal supporting membrane helices (see below). ATPase3 also has the phosphorylation site consensus sequence of FDKTGT[L/I]T, as is typical of P5-ATPases [2]. However, in core D, the substrate-binding kink is typically defined by a PPxxP motif in P5-ATPases [1], and the *Plasmodium* ATPase3 has a PxxxP motif, as commonly found in type-P2A-, type-P2B-, or type-P3-ATPases. Notwithstanding this difference in the substrate-binding kink, ATPase3 is clearly a type-P5-ATPase and shows the most similarity to type-5B, as previously noted [3]. Previous phylogenetic analysis of ATPase1 and ATPase3 also indicated that both are type-P5-ATPases [5].

The unrooted phylogenetic tree generated from complete sequences of *Plasmodium* ATPase3 is characterized by long branches emanating from a central region and radiation into species near the ends of the branches (Figure 1). These branches form seven clades

that somewhat correspond to host species. The seven clades are (1) the subgenus *Laverania*, which are parasites of the great apes, including the human pathogen *P. falciparum*; (2) the vivax-like clade consisting of parasites that infect non-human primates primarily from Asia; (3) the rodent parasite clade; (4) the avian parasite clade; (5) the ovale clade, consisting of *P. ovale curtisi* and *P. ovale wallikeri*; (6) the malariae clade consisting of *P. malariae* and *P. brasilianum*; and (7) the *Hepatocystis* clade, with just a single available sequence thus far. ATPase3 of *P. malariae* and *P. brasilianum* have identical sequences, which is consistent with the high degree of sequence identity previously reported between these two species [13], and thus, *P. brasilianum* ATPase3 was not investigated further.

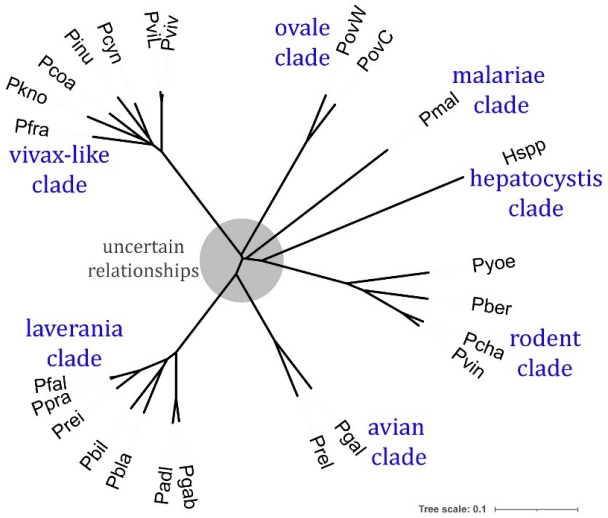


Figure 1. Phylogenetic tree of Atpase3 from *Plasmodium* species. The ClustalW alignment from Supplemental Figure S2 was used to generate an unrooted phylogenetic tree and visualized with interactive Tree of Life [14]. Seven branches, which define seven clades (blue labels), emanate from a central region of uncertain phylogenetic relationships (gray circle). The abbreviations of species names are found in Supplemental Table S1.

The seven clades all emerge from near the center of the unrooted tree, and the phylogenetic relationships between the clades are not clear. This tree structure is observed in many *Plasmodium* genes and reflects the current understanding of *Plasmodium* phylogeny [15]. In the regions corresponding to the transmembrane helices and the eight cores, there is a high level of sequence homology between all the species. However, there are also large regions of variable sequence, which tends to be conserved between the species within a clade but not conserved between the clades (Supplemental Figure S2). This ATPase3 structure of highly conserved domains interrupted by variable sequences that is clade-specific possibly accounts for the tree topology consisting of long branches between the clades and shorter branches within a clade.

ATPase1 and ATPase3 exhibit substantial homology in the eight core regions and the transmembrane helices, and the variations between ATPase1 and ATPase3 are primarily in the variable regions. In addition, the introns of ATPase1 and ATPase3 are located in almost the same position near or within supporting transmembrane helix 3. As found in ATPase1, there are four large variable regions in ATPase3, which were previously designated as inserts [5]. The variable regions of ATPase3, as for ATPase1, are characterized by a large amount of low-complexity sequence, including tandem repeats. Also, as observed in ATPase1, the variable regions are conserved between species within a clade, and sequence divergence within a clade is often associated with tandem repeat variations (Supplemental Figure S2). Within variable region 4, there is a short stretch of approximately

13 amino acids that is conserved in *Plasmodium* species but missing in *Hepatocystis*. The four variable regions are essentially in the same positions for both ATPase1 and ATPase3. A notable difference between ATPase1 and ATPase3 is an extended N-terminus in ATPase3. Overall, the structure of these two genes is very similar, and ATPase1 and ATPase3 are almost certainly related to each other, and ATPase1 likely arose from a duplicative transposition of ATPase3 (see below).

2.3. Non-Plasmodium Orthologues

PfATPase3 and PfATPase1 were used as query sequences in BLAST searches of the NCBI non-redundant sequences, excluding *Plasmodium* and *Hepatocystis* sequences, to identify potential orthologues in other species. To provide a frame of reference, the same BLAST searches were also carried out including only haemosporidian sequences. The highest-scoring ATPase3 hits based on E-values were from *Babesia* and *Theileria* (Table 1). The next highest-scoring hits after the piroplasmids were from the coccidians. There was also a rather high-scoring hit from the colpodellid *Vitrella brassicaformis*. Colpodellids are predatory apicomplexans found at the base of the apicomplexan clade [16]. There was a single hit from *Cryptosporidium*, but its E-value was substantially less than the other apicomplexans, and further investigation of this sequence revealed probable and substantial sequence errors, and it was not investigated further. The high scores among apicomplexan sequences and the lower scores from other species suggest that there may be ATPase3 orthologues in the Apicomplexa, especially the piroplasmids, which are a sister group to the haemosporidians [17]. Complete sequences from these apicomplexans were obtained (Supplemental Table S1) and used in additional phylogenetic analyses (see below).

Table 1. Results of BLAST searches using ATPase3 or ATPase1 as queries in a search for non-Plasmodium orthologs.

Taxonomic Group	PfATPase3	PfATPase1
Piroplasmids	$5 \times 10^{-70} – 2 \times 10^{-101}$	9×10^{-39}
Coccidians	$8 \times 10^{-62} – 6 \times 10^{-64}$	5×10^{-33}
Vitrella brassicaformis (colpodellid)	1×10^{-87}	5×10^{-40}
Cryptosporidia	4×10^{-48}	$2 \times 10^{-30} – 1 \times 10^{-30}$
Ciliates	ND	$6 \times 10^{-31} – 3 \times 10^{-34}$
Oomycetes (stramenopile)	$3 \times 10^{-48} – 1 \times 10^{-53}$	ND
Fungi	$4 \times 10^{-48} – 2 \times 10^{-52}$	$3 \times 10^{-30} – 8 \times 10^{-35}$
Bee (insect)	ND	2×10^{-42}
Anaeramoeba ignava (metamonad)	ND	$4 \times 10^{-32} – 7 \times 10^{-34}$
Haemosporidian ATPase3 sequences	$3 \times 10^{-89} – 0.0$	$3 \times 10^{-18} – 6 \times 10^{-22}$
Haemosporidian ATPase1 sequences	$6 \times 10^{-19} – 9 \times 10^{-23}$	$6 \times 10^{-124} – 0.0$

The range of E-values from a BLAST search excluding haemosporidian sequences using PfATPase3 or PfATPase1 as queries is shown for the indicated group or species. A single E-value means that there was a single hit for that taxonomic group. ND = none detected in the top 100 hits for PfATPase3 or the top 50 hits for PfATPase1. No common subjects (i.e., hits) were identified between the top 100 hits using PfATPase3 as a query and the top 50 hits using PfATPase1 as a query. The ATPase3 and ATPase1 rows are the results of BLAST searches in which only haemosporidian sequences were included.

ATPase3 queries also detected several sequences from oomycetes and fungi (Table 1). Oomycetes are stramenopiles, a sister group to the alveolates, which, with the rhizarians, form the SAR clade, whereas fungi are opisthokonts [18,19]. The phylogenetic relationship between SAR and Opisthokonta is unknown, and this relationship likely

dates back to the base of the eukaryote tree. The similar and overlapping E-values of the oomycetes and fungi and their much lower values than the apicomplexan sequences suggest that none of these sequences are true orthologues. Furthermore, the lack of hits from ciliates and dinoflagellates also suggests that the oomycetes sequences are not orthologues of ATPase3. Apicomplexans, dinoflagellates, and ciliates form the alveolates, with apicomplexans and dinoflagellates forming sister groups combined in the myzozoans [16]. Therefore, the lack of apparent ATPase3 orthologues in the ciliates and dinoflagellates would necessitate at least two separate losses of this gene if there are true ATPase3 orthologues in the oomycetes. Thus, ATPase3 is likely restricted to apicomplexans.

The E-values of the highest-scoring hits in non-haemosporidian species using ATPase1 as the query were substantially lower than those hits of ATPase3, and there were no hits shared between ATPase1 and ATPase3. The majority of the ATPase1 hits were from fungi, and the highest-scoring hit was from an insect (Table 1). There were a few ATPase1 hits within the Apicomplexa, but they were all different proteins than those identified by ATPase3, and these E-values were substantially lower. These included a hit from *V. brassicaformis* (Acc. No. CEL93221.1), a single coccidian sequence from *Besnoitia besnoiti* (Acc. No. XP_029221545.1), and a hit from *Cardiosporidium cionae* (Acc. No. KAF8819569.1). *C. cionae* is a parasite of sea squirts that may be related to piroplasmids [20]. The *C. cionae* hit is a partial sequence corresponding to the C-terminal region which consists primarily of supporting transmembrane helices. Without sequence data from the core ATPase sequences, it is not possible to determine if this *C. cionae* sequence is related to ATPase1, and therefore, it was not investigated further. Likewise, further investigation into the other apicomplexan sequences and representative sequences from the ciliates suggests that these sequences are possibly P5-ATPases but are probably not orthologues of ATPase1. In particular, many of these sequences have extended N-terminal sequences not seen in ATPase1. Thus, no clear orthologues of ATPase1 were detected outside of *Plasmodium* and *Haemoproteus*.

2.4. Phylogeny of ATPase1 and ATPase3

The complete ATPase1 and ATPase3 sequences were aligned and used to generate phylogenetic trees using *Vitrella brassicaformis* ATPase3 as the outgroup. Overall trees generated by Maximum Likelihood, Neighbor Joining, or Minimum Evolution had similar topology with minor differences. Differences between these methods were primarily in the placement of *Hepatocystis* and *Babesia microti* and the exact branching order of the ATPase3 sequences within *Plasmodium*. To minimize the effect of the variable regions, the variable regions were removed from the sequences, and the resulting trees were similar to trees generated with complete sequences with similar minor differences. The branching order exhibited by ATPase3 from the colpodellids, coccidians, piroplasmids, and malaria parasites (Figure 2) exhibited the expected and well-known phylogeny of the Apicomplexa [16]. This further corroborates that these apicomplexan ATPase3 sequences are true orthologues.

Plasmodium ATPase3 sequences did not branch in the expected order, and the order varied slightly depending on the method used to generate the trees and whether complete or conserved sequences were used. However, this is of little consequence since molecular phylogeny based on single genes is rather unreliable. ATPase1 sequences consistently formed a sister clade with the *Plasmodium* ATPase3 sequences regardless of the methods or sequences used. The ATPase1 branch is also rather long compared to the other branches, thus implying that an accelerated rate of evolution occurred after the duplication. A faster rate of evolution following gene duplication was also reported for acyl-CoA synthetase paralogues of *P. falciparum* [21]. This accelerated divergence between ATPase1 and ATPase3 likely precludes the ability to place the ATPase1 clade within the *Plasmodium*

ATPase3 clade, as would be expected for these two paralogues. Nonetheless, this phylogeny demonstrates that the duplication and transposition of ATPase3 leading to the formation of ATPase1 likely occurred early in the evolution of malaria parasites.

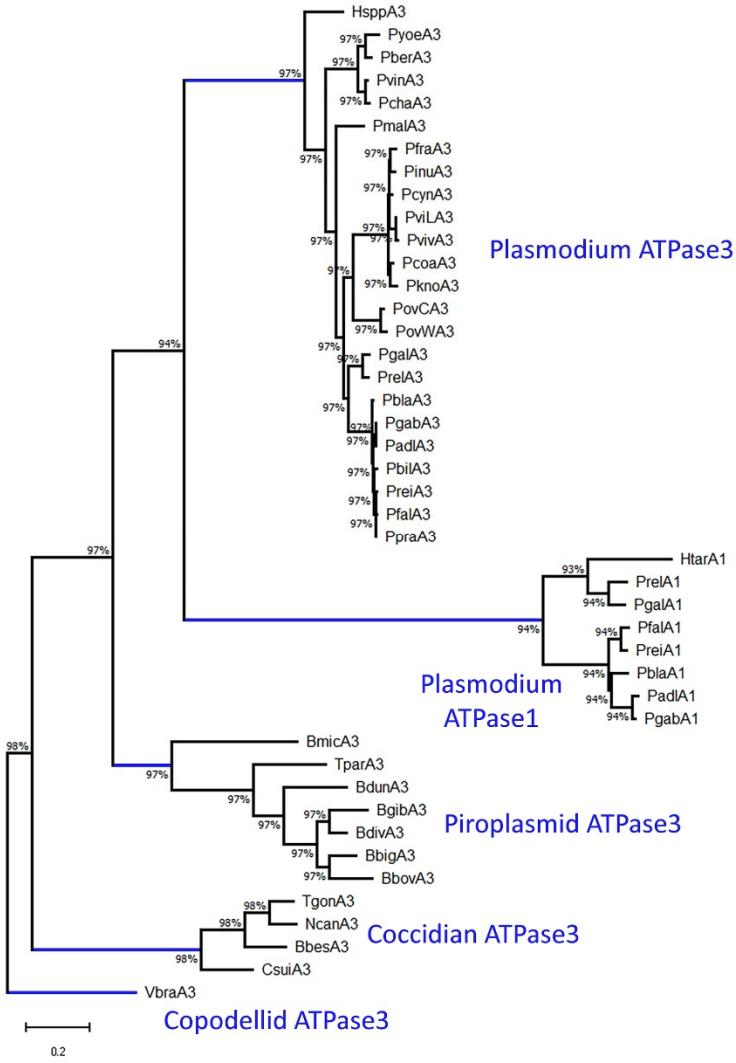


Figure 2. Phylogeny of ATPase1 and ATPase3 from Apicomplexa. The variable regions from the ATPase1 and ATPase3 sequences were removed, and the remaining conserved sequences were aligned with MUSCLE within MEGA XI [22]. A phylogenetic tree using Maximum Likelihood was generated from the alignment. Descriptions of the species and the abbreviations are found in Supplemental Table S1, and bootstrap values are shown at the nodes. The five major clades (denoted in blue) are ATPase3 from colpodellids, ATPase3 from coccidians, ATPase3 from piroplasmids, ATPase3 from *Plasmodium* including *Hepatocystis*, and ATPase1 from *Plasmodium* including *Haemoproteus*.

2.5. Synteny of ATPase1 and ATPase3

The duplication of ATPase3 was further investigated by the characterization of synteny. ATPase1 and ATPase3 are located in two different syntenic blocks on chromosome 5 in the *Laverania* species and on chromosome 10 in the avian parasites (Figure 3). The genes flanking ATPase3 are the same in all *Plasmodium* species and are the AN1-type zinc-finger protein and a conserved *Plasmodium* protein of unknown function. These same

two genes also flank the ATPase3 of *Hepatocysitis* on contig CABPSV020000075.1. This same ATPase3 syntenic block is found on chromosome 10 of the vivax-like parasites and *P. ovale*, chromosome 6 in *P. malariae*, and chromosome 11 of the rodent parasites. These chromosomal designations are consistent with the known synteny of *Plasmodium* chromosomes [23]. The synteny of ATPase3 from *Haemoproteus tartakovskyi* could not be determined since only a partial genomic sequence is available. No synteny of ATPase3 from the other apicomplexans with ATPase3 of *Plasmodium* was found. This is expected since a low level of synteny—which is restricted to microregions—between *Babesia* and *Plasmodium* was previously reported [24].

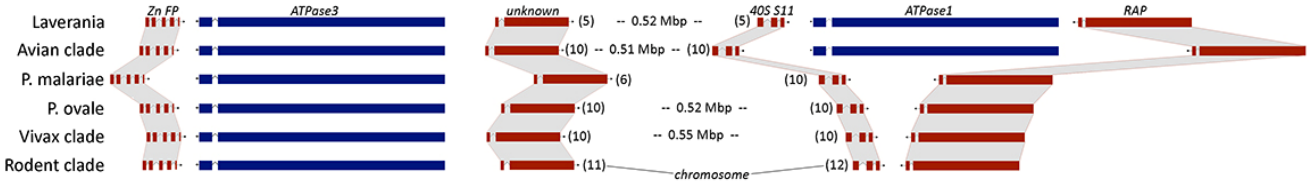


Figure 3. Synteny of ATPase1 and ATPase3. Synteny of the ATPase1 and ATPase3 genes between *Plasmodium* species was originally analyzed with JBrowse [25] embedded within PlasmoDB [26]. ATPase1 and ATPase3 were aligned (blue), and the relative positions of the flanking sequences are shown in orange. The flanking sequences are an AN1-type zinc-finger protein (ZnFP), a conserved unknown protein, the 40S ribosomal protein S11, and a putative protein with a RAP domain. The numbers in parentheses are the assigned chromosomes. In the parasite clades that have ATPase1 and ATPase3 on the same chromosome, the average distance between the genes in indicated in million base pairs (Mbp).

ATPase1 is also found on chromosome 5 of *Laverania* species and chromosome 10 of avian malaria parasites and is immediately flanked by the 40S ribosomal protein S11 and a putative protein with a RAP (RNA-binding domain abundant in Apicomplexans) domain (Figure 3). As in *Laverania* and avian *Plasmodium* species, 40S ribosomal protein S11 is found on the 3′-side of the ATPase1 gene and in the opposite orientation in *H. tartakovskyi* (Appendix A). The gene on the 5′-side of *H. tartakovskyi* ATPase1 could not be determined due to limited available DNA sequence on this side of the gene. This same syntenic block between the S11 ribosomal protein and RAP genes is found in the *Plasmodium* species without ATPase1 and is located on chromosome 10 of vivax-like parasites, *P. ovale*, and *P. malariae* and located on chromosome 12 in the rodent parasite clade. These results are consistent with the known synteny of *Plasmodium* chromosomes [23]. Chromosome 5 of the *Laverania* species approximately corresponds to chromosome 10 of *P. relictum*, members of the vivax-like clade, and *P. ovale*. In *P. malariae*, a portion of chromosome 10 has broken off to form chromosome 6. In the rodent malaria clade, approximately half of the chromosome 5/10 syntenic block is associated with chromosome 11, and the other half of this block is associated with chromosome 12. In the clades in which these two syntenic blocks are located on the same chromosome, the distance between ATPase1 and ATPase3 is approximately a half-million base pairs.

2.6. Predicted Three-Dimensional Structures of ATPase1 and ATPase3

Select sequences of ATPase1 and ATPase3 genes that represent distinct apicomplexan and *Plasmodium* clades were used to search protein databases for proteins with similar 3-dimensional structures. Numerous proteins were identified with 100% confidence of homology, and the coverage in the top-scoring templates ranged from 30 to 56%. The regions that were not modeled were primarily in the N-terminal extension of ATPase3 and the variable regions of both paralogues (Supplemental Figure S3). The top-scoring

templates were heavily shared among the various query sequences (Table 2), and most of the identified templates were designated as P5B-ATPases. Overall, the predicted structures of ATPase1 and ATPase3 are quite similar regardless of species or template. One exception is that the top-scoring template of ATPase3 from *P. falciparum* and *Hepatocystis* was a sodium–potassium pump from *Sus scrofa* (wild boar). Overall, this template was a rather low-ranked template among the other query sequences. This template produced predicted structures that were quite different from P5B-ATPases and was not investigated further.

Table 2. Ranks of common and top-scoring templates identified by Phyre® with ATPase1 (A1) and ATPase3 (A3) sequences.

Species (Queries)	7op1	6xms	8ier	7m5x	7n78	7fjp	3b8c
Plasmodium falciparum (A1)	1	7	5	4	6	3	2
Plasmodium relictum (A1)	1	6	4	2	3	5	31
Haemoproteus tartakovskyi (A1)	1	7	3	2	5	6	40
Plasmodium falciparum (A3)	2	6	5	4	7	8	1
Plasmodium relictum (A3)	1	3	4	6	2	5	10
Plasmodium vivax (A3)	5	6	2	4	1	3	26
Plasmodium malariae (A3)	1	2	3	4	5	6	18
Plasmodium ovale (A3)	2	6	1	5	3	4	32
Plasmodium yoelii (A3)	1	2	3	4	5	6	7
Hepatocystis spp (A3)	3	2	10	4	7	6	1
Vitrella brassicaformis (A3)	5	6	2	4	3	1	13
Theileria parvum (A3)	1	6	3	5	2	4	19
Babesia divergens (A3)	2	6	5	3	1	4	33
Toxoplasma gondii (A3)	1	6	3	4	2	5	7
Cystoisospora suis (A3)	3	6	5	4	2	1	24

Sequences from the indicated species were used as queries and analyzed by Phyre®, . Shown are the ranks of common templates (denoted with PDB IDs in column headers) for each of the queries. These templates are Ypk9 from *Chaetomium thermophilum* (7op1), ATP13A2 of *Homo sapiens* (8ier, 7m5x, 7n78, or 7fjp), a P5A-ATPase from *Saccharomyces cerevisiae* (6xms), and a sodium–potassium pump from *Sus scrofa* (3b8c).

Supplemental Figure S4a shows models generated from the top-scoring templates—or the second highest-scoring template in the case of ATPase3 from *P. falciparum* and *Hepatocystis*—for each of select sequences. In general, these predicted 3-dimensional structures were all quite similar, and the major domains of P-type ATPases are readily identified. The most notable differences between the species are in the region that corresponds to the N-terminal membrane helix loop (nMHL). Overall, the 7m5x template of ATP13A2 generated the best consensus models for all the sequences (Supplemental Figure S4b). However, this template did exclude the N-terminal membrane loop and part of the N-terminal domain from the ATPase3 proteins of *Laverania*, *Toxoplasma*, and *Cystoisospora*. None of the templates included the N-terminal membrane helix loops in the *Laverania*. Perhaps the rather long N-terminal extension or a rather large variable region 1 may interfere with the modeling.

Predicted 3-dimensional modeling was used to compare the substrate-binding sites of ATPase1 and ATPase3. In all P-type ATPases, the core transmembrane helices form a channel that faces the luminal (or extracellular) side of the ATPase, and the central kink within cTM4 closes off this channel (Figure 4). The conformation change associated with the phosphorylation–dephosphorylation cycle opens an exit channel to the cytoplasm at this central kink. No overt differences between the structures of ATPase1 and ATPase3

were observed at the whole-molecule level. ATP13A2 is a particularly well-characterized P5B-ATPase [27], and the predicted structures of the ATPase1 and ATPase3 substrate-binding sites were compared to the known structure of ATP13A2 (PDB ID 7m5x). The substrate-binding site of ATP13A2 is stabilized by interactions between a tyrosine in cTM2, a tyrosine in cTM5, a glutamine in cTM5, and an aspartate in cTM6. An aspartate found at the boundary between cTM1 and cTM2 and another aspartate in cTM4 provide negative charges to interact with the positive charges of polyamines. Negative charges are also seen at the entrance to the luminal channel in the Ypk9 protein of yeasts [28]. ATPase1 is identical to ATP13A2 at these six key residues (Figure 5). Furthermore, the sequence of the central kink is nearly identical between ATPase1 and ATP13A2 and has a consensus sequence of [I/V/L]PPALP.

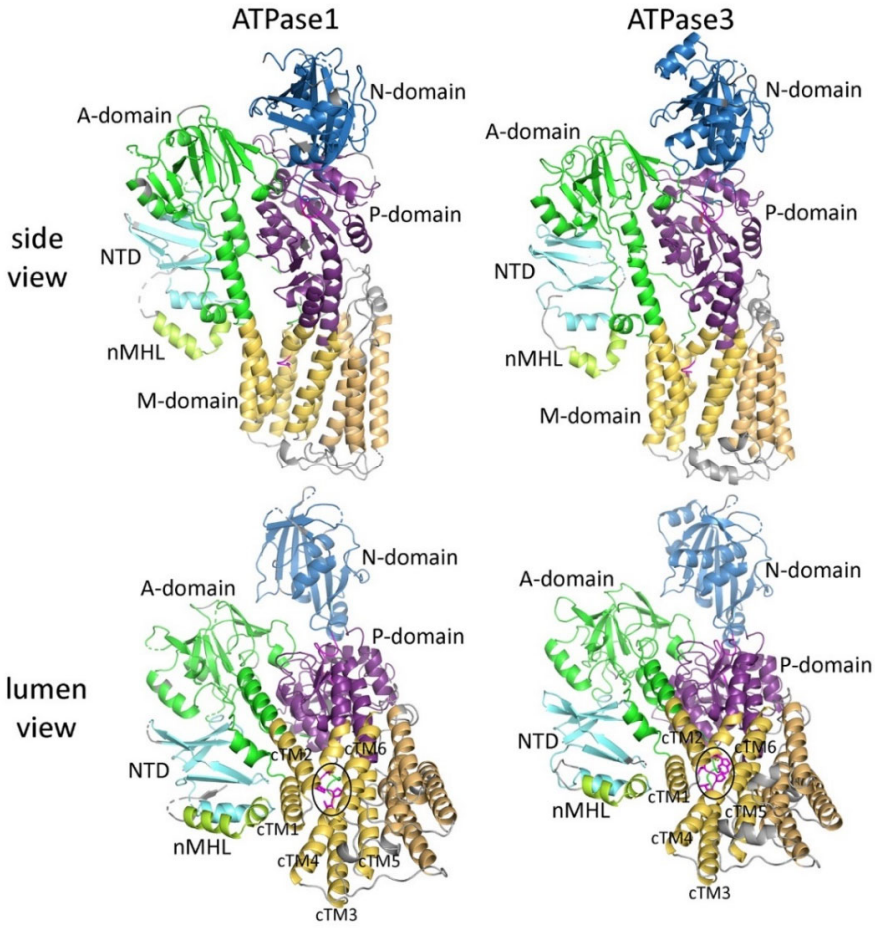


Figure 4. Ribbon diagrams of ATPase1 and ATPase3 from *Plasmodium relictum*. Models were generated by Phyre® and visualized by PyMOL® using the template of a P5B-ATPase from *Homo sapiens* (PDB ID 7m5x). In the top two panels, the cytoplasmic-facing portions of the proteins are up, and the lumen (or extracellular)-facing portions of the proteins are facing down. Major domains are denoted and colored according to the following scheme: nMHL (yellow), cTM (yellow-orange), sTM (light orange), NTD (aquamarine), A-domain (green), N-domain (sky blue), P-domain (violet purple), the kink in cTM4 (magenta), and the phosphorylation motif (magenta). The combined cTM and sTM are denoted as membrane (M) domain. Gray denotes residues not assigned to a domain or positions where residues were removed by the modeling program. In the bottom two panels, the molecules are rotated to show the channel formed by the six core transmembrane helices (cTM#), which opens on the luminal side. The channel is transiently closed by the central kink in cTM4 (magenta side chains) and is circled.

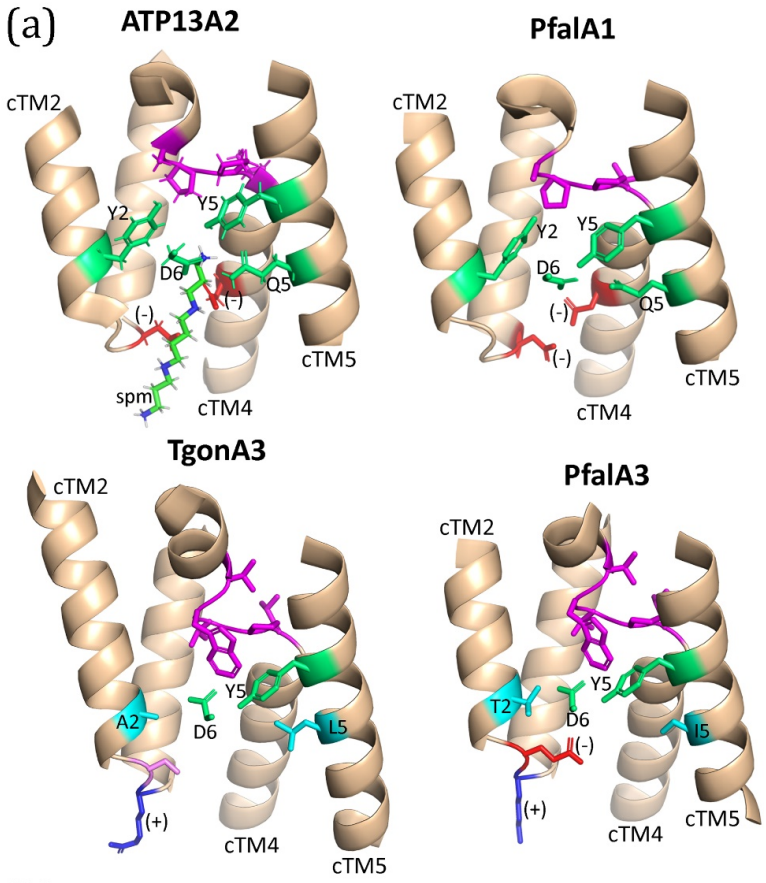


Figure 5. Predicted binding site of ATPase1 and ATPase3. Shown are ribbon-diagram models of the substrate-binding channel (**a**) and alignments of the key residues forming this substrate-binding site (**b**). ATP13A2 (Acc. No. NP_071372.1) was used as a template (PDB ID 7m5x) to generate these structures and includes the bound spermine (spm). The six core transmembrane helices (cTM) form a channel that serves as the substrate binding site, with the kink in the middle of cTM4 (magenta side chains) closing this channel. Three of the cTMs are labeled, and cTM1 is unlabeled in the background. cTM3 and cTM6 are not shown to better reveal the substrate-binding channel. Interactions between a tyrosine in cTM2 (Y2), a tyrosine in cTM5 (Y5), a glutamine in cTM5 (Q5), and an aspartate in cTM6 (D6) in ATP13A2 stabilize the binding channel, and the side chains of these four residues are shown (green). Conservation of this tetrad in the apicomplexans is also colored green, and deviations are colored aqua. Side chains of charged residues are also shown with red denoting negative (−) charges and blue denoting positive (+) charges. A polar replacement of a negative charge is colored pink. Two aspartate residues in ATP13A2 participate in binding the spermine substrate. The three basic types of substrate-binding sites are shown, and the substrate-binding sites of all

fifteen select apicomplexan sequences are shown in Supplemental Figure S4c. The alignment shows the key residues making up the substrate-binding site, with membrane helices shaded in gray and the four residues making up the tetrad denoted with T. Conserved tetrad residues are highlighted in green, and deviations are highlighted in aqua. Dark-green shading denotes polar residues next to a polar tetrad residue. Negatively charged residues (−) are highlighted in red and positive residues (+) in blue. Polar replacements of the glutamate residues (E) are shaded in purple. The abbreviations of species names are found in Supplemental Table S1.

ATPase3 deviates from ATPase1 in the key residues that make up the substrate-binding site (Figure 5). In the stabilization tetrad, the tyrosine in cTM5 and the aspartate in cTM6 are conserved, allowing for conservative replacements of phenylalanine for tyrosine in cTM5 and glutamate for aspartate in cTM6 in the piroplasmids. The tyrosine in cTM2 is replaced by either serine, threonine, or alanine, and the polar glutamine in cTM5 is replaced with a hydrophobic residue (isoleucine or leucine). However, there is a polar residue (serine or threonine) adjacent to this position (except for the piroplasmids). The two aspartates that bind to polyamines are also absent in ATPase 3 and are replaced with either tryptophan or glycine. However, a negatively charged glutamate residue is found in the loop between cTM1 and cTM2 in ATPase3 from *Plasmodium*. In the other apicomplexan species, this position is occupied by a polar group (serine or asparagine). A notable difference between ATPase1 and ATPase3 is a positively charged residue (lysine or arginine) in the loop between cTM1 and cTM2. In addition, the kink in ATPase3 deviates from ATPase1 with the consensus sequence of [L/V/A]P[V/I]WT[P/S]. These results suggest that ATPase1 and ATPase3 exhibit different substrate specificities.

3. Discussion

ATPase1 and ATPase3 are two related P-type ATPases that were initially identified in the malaria parasite [4–6]. Sequence alignments with the eight P-type ATPase cores (A–H), BLAST searches, and three-dimensional structure predictions are consistent with both of these ATPases being P5B-ATPases, as previously reported [3]. Both ATPase1 and ATPase3 genes consist of two exons, and the introns in both genes are located in almost exactly the same position. Both genes also contain four major variable regions that are located in the same positions in both paralogues. As with many *Plasmodium* genes [11,12], these variable regions tend to be low-complexity sequences and often contain tandem repeats of variable length. Differences between ATPase1 and ATPase3 include an extended and variable N-terminus in ATPase3, the predicted substrate-binding site, and species distribution. The presence of N-terminal extensions has been reported in other P5B-ATPases, and these extensions have been speculated to perform an auto-inhibitory role by folding back into the phosphorylation site [28]. The high degree of variability among the apicomplexan sequences, however, does question the functional significance of these N-terminal extensions.

P5-ATPases are the least characterized among the P-type ATPases. Two of the better-characterized P5B-ATPases are Ypk9 of yeasts [28] and ATP13A2 of humans [27]. Both of these P5-ATPases are likely polyamine transporters that move polyamines from a lysosome-like compartment to the cytoplasm. However, it has also been proposed the Ypk9 plays a role in sequestration of heavy metals [29]. The primary lysosome-like compartment in the malaria parasite is the digestive vacuole [30]. The digestive vacuole of the malaria parasite has been extensively studied, and a thorough proteomic analysis did not detect any P-type ATPases [31], suggesting that neither of these *Plasmodium* paralogues are in the digestive vacuole. Previous immunofluorescence studies have suggested that ATPase1 and ATPase3 are localized to an intracellular compartment [4–6]. ATPase1 was

reported to also localize in the region of the parasite plasma membrane [4]. However, inspection of the published micrographs reveals a diffuse staining throughout the cytoplasm of the parasite that is similar to those published for ATPase3, and thus, ATPase1 does not appear to be plasma-membrane-associated. The previous studies were all carried out with acetone-fixed thin blood smears, which limits the resolution and ability to identify subcellular compartments. Therefore, without additional studies, it is not possible to definitively assign ATPase1 nor ATPase3 to a specific cellular compartment.

ATPase1 most likely arose from a duplication and transposition of ATPase3 in the early evolution of malaria parasites since ATPase3 orthologues are found throughout the Apicomplexa, and thus far, ATPase1 is only found in *Haemoproteus*, avian *Plasmodium* species, and *Laverania*. The observation that ATPase1 has been retained in these three distinct taxa suggests that ATPase1 has a functional role that has been maintained. Furthermore, the predicted differences in the substrate-binding sites of ATPase1 and ATPase3 imply distinct functions for these paralogues. Thus far, polyamines have been the most investigated substrates of P5B-ATPases, and ATPase1 has a predicted substrate-binding site that is very similar to the polyamine-binding sites of ATP13A2 and Ypk9 (Figure 5). ATPase3 differs from ATPase1 in the predicted substrate-binding site, and in particular, there are fewer negatively charged side chains in the binding site, which would function to complex the positive amine groups of polyamines. In addition, the predicted binding site of ATPase3 from *Plasmodium* is slightly different than the predicted binding site of ATPase3 from the other apicomplexans (Figure 5).

Malaria parasites have the three core enzymes of polyamine biosynthesis, which include ornithine decarboxylase, adenosylmethionine decarboxylase, and spermidine synthase [32]. *Plasmodium* is unique in that ornithine decarboxylase and adenosylmethionine decarboxylase are combined into a bifunctional enzyme. A search of PlasmoDB revealed that the bifunctional enzyme and spermine synthetase are found in all *Plasmodium* species as well as *Hepatocystis* and *Haemoproteus*, and these enzymes are highly conserved. These enzymes are presumably located in the cytoplasm since polyamines are needed for the modification of a translation elongation factor. In addition, polyamines are also taken up by the parasite directly into the cytoplasm via an electrogenic mechanism that depends on the parasite's negative membrane potential [33]. Therefore, there is no need for a transporter to move polyamines to the parasite cytoplasm. Furthermore, it is likely that polyamine metabolism does not differ among the malaria parasites, and a distinct polyamine transporter in some species is probably not needed. So, probably neither ATPase1 nor ATPase3 are polyamine transporters.

In the initial reports, ATPase3 was expressed predominantly in gametocytes as determined by immunofluorescence [5,6], and ATPase1 was only investigated in late ring and early trophozoite stages [4]. Analysis of mRNA expression for ATPase3 has maximal expression in either the late ring stage [7] or the trophozoite and schizont stages, with no expression in the gametocyte stages [5]. Examination of the transcriptomic studies available through PlasmoDB [26] also reveals inconsistencies in the stage specificity of expression for both paralogues. Both genes appear to be expressed in all stages of the life cycle, including the vector stages. Therefore, at this time, nothing definitive can be said about the stage specificity of expression of either paralogue. Clearly, more work is needed on the subcellular location, stage specificity of expression, and activity of these two paralogues.

Phylogenetic Implications of ATPase1

The presence of ATPase1 in *Laverania*, avian *Plasmodium* species, and *Haemoproteus* and its exclusion in all other Apicomplexa, including other mammalian *Plasmodium* spe-

cies and *Hepatocystis*, strongly suggest that *Laverania* is more closely related to avian malaria parasites including *Haemoproteus* than to the other mammalian malaria parasites. The most parsimonious explanation is that ATPase3 underwent a duplicative transposition during the early evolution of the malaria parasite only in a lineage containing *Laverania*, avian *Plasmodium* species, and *Haemoproteus*. The combined phylogeny of ATPase1 and ATPase3 shows *Plasmodium* ATPase3 and ATPase1 diverging at the base of the *Plasmodium* clade (Figure 4), which is consistent with the appearance of ATPase1 early in the evolution of malaria parasites. In contrast, several genome-level molecular phylogenetic studies suggest that all malaria parasites of mammals form a single clade, with *Laverania* placed as a sister group to the other mammalian parasites, and the avian malaria parasites form a sister group to an exclusive mammalian parasite branch [34–36]. To make this molecular phylogeny congruent with the ATPase1 data would require subsequent losses of the ATPase1 gene in at least one mammalian parasite lineage in the appropriate time frame and therefore is less parsimonious.

The availability of complete genome sequences has strengthened molecular phylogeny. For example, concatenated sequences from many genes provide more resolution and stronger statistical arguments. Despite the robust statistics, the methodology is nonetheless a summation of single genes, which tends to be unreliable. Indeed, one study using concatenated genes found a well-supported clade consisting of *Laverania* and avian malaria parasites [37]. Molecular phylogeny should be complemented with other approaches, such as taxon-specific characters or gene synteny. Specific genes, such as ATPase1 described herein, could be considered as taxon-specific characters. The use of ATPase1 as a character changes the possible phylogeny of *Plasmodium* and *Haemoproteus* within the Haemosporidia (Figure 6).

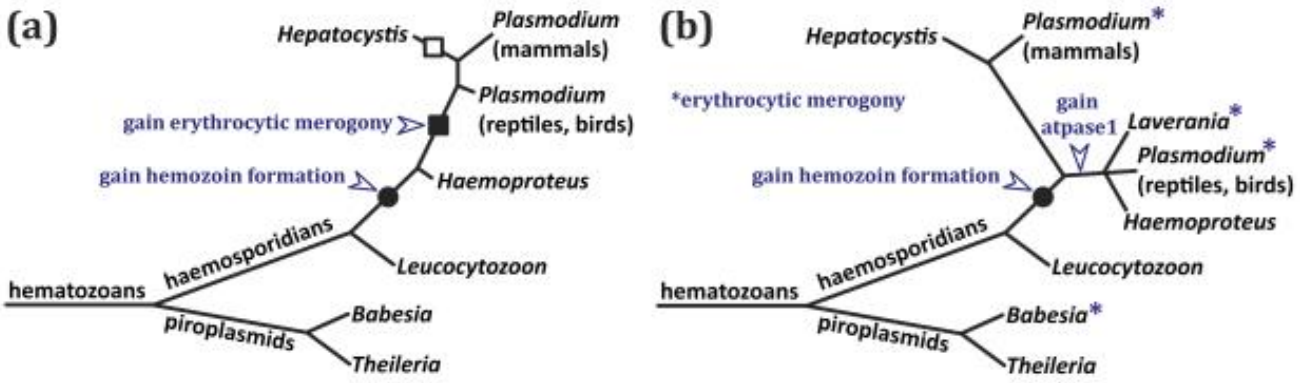


Figure 6. Possible change in the phylogeny of haemosporidians associated with ATPase1. The panel on the left (**a**) shows possible branching order of major genera [16,17], as modified from Figure 3 of Wiser (2024) [30] with permission. The panel on the right (**b**) shows how the phylogeny changes with the gain of ATPase1 (arrowhead). Branch lengths do not depict evolutionary distances, and branches do not depict the true complexity of Haemosporidia. The gain of hemozoin formation is denoted with a filled circle, and gain or loss of erythrocytic merogony is denoted with a filled or open square (**a**), and asterisks (*) denote phyla that exhibit erythrocytic merogony (**b**).

Haemosporidia is a diverse group of blood parasites in reptiles, birds, and mammals that are transmitted by blood-feeding dipterans. The phylogeny in this group is complicated and uncertain [38]. For example, among the three major clades within Haemosporidia, represented by *Leucocytozoon*, *Haemoproteus*, and *Plasmodium*, two (*Haemoproteus* and *Plasmodium*) are considered polyphyletic. The two characters used for the morphological phylogeny of haemosporidians are hemozoin and erythrocytic merogony. As previously

noted [34], erythrocytic merogony is an unreliable character for phylogeny, as it is sporadically observed throughout the haematozoans (Figure 6b). Gains and losses of erythrocytic merogony have occurred several times within the haemosporidians [39], and genes associated with erythrocytic merogony have been lost in *Hepatocystis* [40]. Furthermore, since cell division is generally regulated, merogony could be a phenotypic trait. For example, dormant blood-stage forms that do not replicate have been described in *Plasmodium* [41,42].

Clearly, more sampling of avian and reptilian *Plasmodium*, *Haemoproteus*, and *Leucocytozoon* species is needed to resolve the phylogeny of Haemosporidia. Going forward, it would be worthwhile to investigate the distribution of ATPase1 among these taxa. In addition, there may be other genes that could serve as reliable taxonomic characters. For example, Bohme et al. (2018) identified 15 genes, including ATPase1, that appear to be restricted to *Laverania* and avian *Plasmodium* species [35]. Likewise, two orthologous clusters are shared between *P. falciparum* and *H. tartakovskyi* that are absent in other mammalian *Plasmodium* species [34]. One is an unknown protein (PF3D7_1004100), and the other is reticulocyte-binding protein homolog-1 (Pf3D7_0402300). A search for orthologues to these two proteins was carried out since numerous whole genomes have been added to the databases in the last eight years, including several *Laverania* species, *P. relictum,* and *P. gallinaceum*. The unknown protein has syntenic orthologues in the other *Laverania*, *P. relictum,* and *P. gallinaceum* (Ortholog Group: OG6_533697), and the reticulocyte-binding protein homolog-1 has syntenic orthologues only in *Laverania*. Any of these genes may also prove to be reliable taxonomic characters and merit some further investigation.

4. Materials and Methods

4.1. BLAST Searches

ATPase1 (Genebank Acc. No. X65738) and ATPase3 (Genebank Acc. No. X65740) were used as queries in BLASTP searches [43] of non-redundant protein sequences at NCBI. Initially, BLAST searches were restricted to members of the Haemosporidia to identify all ATPase1 and ATPase3 orthologues. Alignments associated with the BLAST results were used to eliminate partial sequences and sequences with obvious errors. Subsequently, the newly identified sequences were used as queries in BLAST searches to ensure that no complete sequences of ATPase1 and ATPase3 in the Haemosporidia were missed. In cases where complete sequences were available from multiple strains of a single species, the reference strain for that species was chosen as much as possible.

A partial genome sequence of *Haemoproteus tartakovskyi* is available [34], which is not part of the non-redundant database. Therefore, a TBLASTN search of those contigs was performed using ATPase1 and ATPase3 sequences from *Plasmodium relictum*. A complete ATPase1 sequence and a partial ATPase3 sequence were identified (Appendix A).

To identify orthologues in non-haemosporidian species, the same strategy of BLAST searches was performed, except haemosporidian species were excluded from the search. Supplemental Table S1 contains a description of all the sequences analyzed in this study.

4.2. Alignments and Tree Building

ClustalW at https://www.genome.jp/tools-bin/clustalw or alignment programs within Mega XI [22] were used to generate alignments. Alignments were adjusted manually to accommodate variable regions and core regions of P-type ATPases. Pairwise alignments with known P5-ATPase sequences were used to assist in identifying these cores. Separate alignments with the various *Plasmodium* clades were also used to optimize the alignments within variable regions. Potential transmembrane helices and membrane domains were identified with CCTOP [44].

Phylogenetic trees were built using Interactive Tree of Life [14] or Mega XI [22].

4.3. Three-Dimensional Structure Predictions

The complete sequences were used to search for conserved three-dimensional structure using Phyre® (https://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index, accessed on 10 January 2025) [45]. Structures were generated using PyMOL® Molecular Graphics System, Version 3.0.4 (Schrödinger, LLC).

5. Conclusions

ATPase1 and ATPase3 are paralogous type-P5B-ATPases of malaria parasites. ATPase3 is found throughout the Apicomplexa, and thus far, ATPase1 is only found in *Laverania*, avian *Plasmodium* species, and *Haemoproteus*. ATPase1 most likely arose from a duplication and transposition of ATPase3 early in the evolution of the malaria parasite and may be a useful taxonomic/phylogenetic character to decipher the phylogeny of Haemosporidia. Both ATPase1 and ATPase3 are likely associated with an intracellular compartment within the blood-stage parasite. The exact compartment, stage specificity of expression, and substrate specificity of both ATPases still need to be determined.

Supplementary Materials: The following supporting information can be downloaded at www.mdpi.com/10.3390/parasitologia5010006/s1. Supplementary Table S1: ATPase1 and ATPase3 sequences evaluated in this study; Supplementary Figure S1: Annotated alignment of ATPase1 from *Plasmodium* and *Haemoproteus*; Supplementary Figure S2: Annotated alignment of ATPase3 from *Plasmodium* species and *Hepatocystis*; Supplementary Figure S3: Alignment of select ATPase1 and ATPase3 sequences from Apicomplexa; Supplementary S4: Three-dimensional modeling of select ATPase1 and ATPase3 sequences.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Contact the author to receive the data generated by this study.

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A

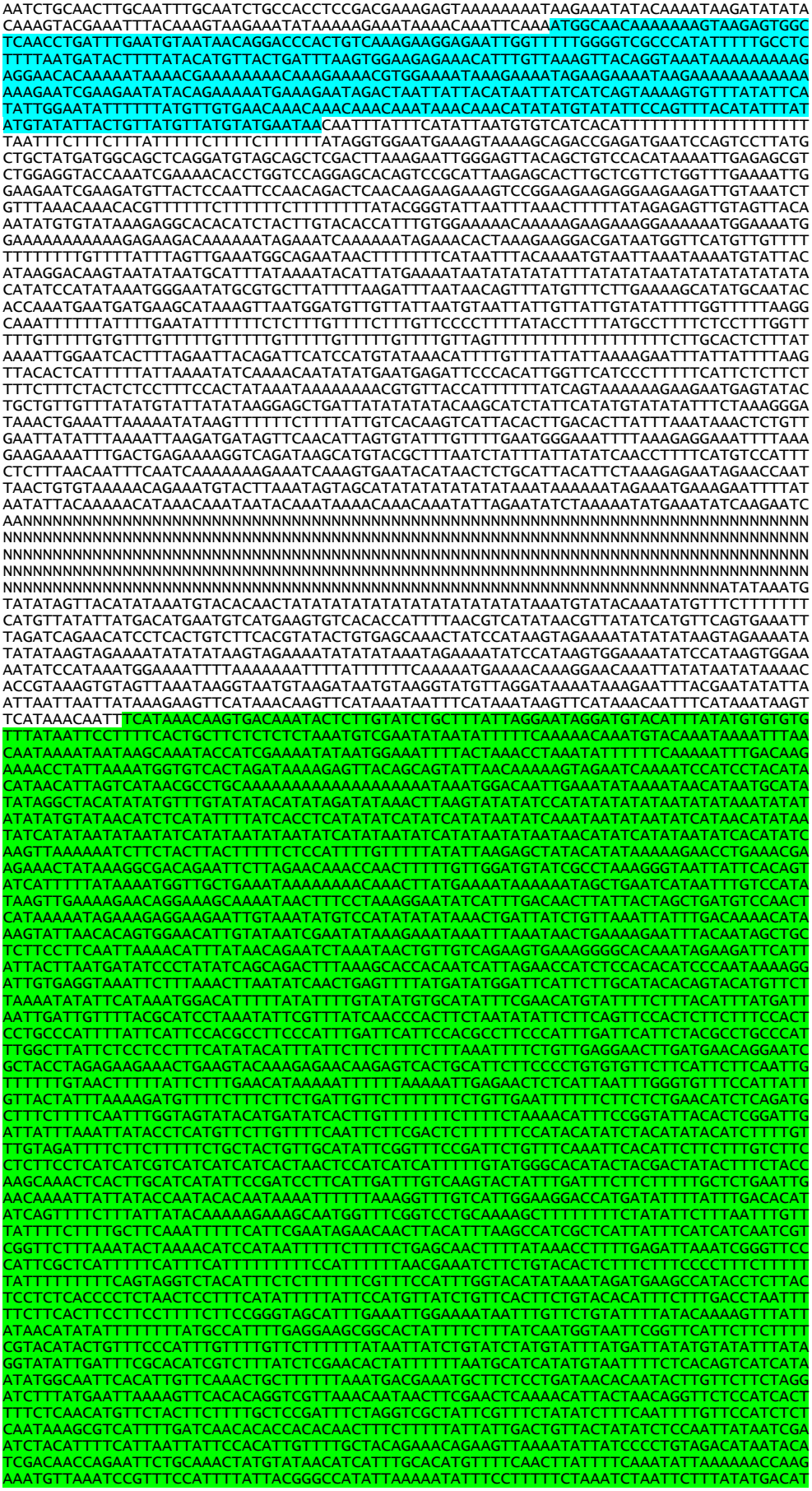
Description of ATPase1 and ATPase3 from Haemoproteus tartakovskyi

The genome of *Haemoproteus tartakovskyi* [34] was searched with ATPase1 and ATPase3 sequences at https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_001625125.1/ (accessed on 10 January 2025). ATPase1 identified a single contig of 18,840 bases (LSRZ01000339.1). The ATPase1 coding sequence started at position 17,290 and ended at position 10,048 (Figure A1). A single intron was located at positions 10,705–10,377 (Figure A2) in the exact same location as other *Plasmodium* ATPase1 genes. As expected, no genes were found on the short 5′-side of the ATPase1 gene. The next gene on the 3′-side of ATPase1 is the 40S ribosomal protein S11 gene, and following that is the putative tRNA pseudouridine synthase gene (Figure A1), which are the same two genes found in *Plasmodium* species.

ATPase3 identified a single contig (LSRZ01001921.1) of 2037 bases (Figure A3). The reverse complement of this sequence encoded a partial protein sequence of ATPase3 that included core domains B–E, all of variable region 2 and most of variable region 3 (Figure A4). There was not a continuous open reading frame in contig LSRZ01001921.1, and Fig-

ure A5 shows possible locations for frame shifts within variable regions 2 and 3. The reading frames were aligned with *P. relictum* ATPase3 to identify possible sites for frameshifts. However, the two frameshifts occurred in variable regions, and it was not possible to unambiguously determine the frameshifts.

>LSRZ01000339.1 Haemoproteus tartakovskyi strain SISKIN1 HtScaffold0339, whole genome shotgun sequence
CATATTTGCTGTGTTTGCTTTGTCCCTCTAATATTCTCTTTATAAAATTTTCTTTGTAGAATTCTTATTGCAG
ATATATATTTATATATATATCTCTCTCTCTCTCTCTCTCTCTCTCTCATATATATATACATATATATATATCTATC
ATTTGTTTTTTTTCATGTTTTATTAATCAACATGTCGACAACCAATCCTACAAATCAGACAACGGAGTGGGACGATTTGCA
AAGAAAGTTTGGAAATTTGCCACCACTAGAAAAGGAAGTAAAAGAAGAAGAATTATATTTAAAAGCTATAGATAAAATTGA
AAATGCCGATATTTTAGAAAACAAAAATTTAAATGAATTAAATCGATTAGAAGAAAATTGTGTCGATGATGAATATTTAAA
AATTATTCAAAAATATAAAAATGATCGATTAAACGAAATAAAAAAAAATAAGGCATTAGAAATATTTGGGGATGTGTATGA
AATAACTAAAGAAAATTTTGTTACTGAAATTAATGATGCTAGTGCATTCAGTCCCTTAAATAAAATAAATGATCAGGAAAA
AGAAAATGATGCAAGAGAAATGAATCAACAAGAATTGGATGAACAAGAAATGAATGAACAAGAAATGAAAAAAAAAAGAAC
AAAAAATGATAATGGAAAAAAAAGGAAAATATAAAAAGTTTGGAACATATGTTATTTTACATCTTTATTCCGAAAATGTTA
TGGCTTGTAAAATTTTAAATAATATTATGAAAGAAGCAGCTAGAAAATATAAATATATAAAATTTACAAAAGGAGTACATA
ATAAAATTATAGAAAATTATCCAGAATCAAAAACTCCAACCATTTTAATATATTATAATGGAATATGTATTCAACAAATAT
GTAATTTATTACTATATATAGGAGGAGGATTATCGAATTTGAATTTAAAATCATTTGAAAAATTGTTAAAAAAATATGATA
TTATAAAAATAAAAAAACAAAAAAAAGGAAATGATGCTCTTAGTGACGATAGTGATGAGGAAGATGAGGAAAATATAAAAA
ATAGAAAAACACAAAAACAATATACTGCTTTTAATATGTTTAATAAAACTTATACAAAAGCAAAAGATGATGATGAAGATA
ATTCAAGTGATAAAACAGAAGAACAAGAAAACTCTACTTCAAAAGGATATGCTTCCTCTTTTTTGGATAATCAGATGAGAA
GATATCGATTTTAAAAAAAACAAAAAAAATAGGATATGAAAAAGGAAAATATAATATAATATAATGAGTTTAATCGTGTAA
TACGTATGTTAAAGGAAAATTAAAAAACAGTTATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTTTATTTTTTTATTTCGTTTTTTTTTTTATTTTATTTGATTAAATAT
ATTAAAATAAAAGATGAGAAAGAAATGTGTAAAGAAATGTATATTTACATATTATATATATAGATATAGATATAGATAGAT
ATTGATATAGATAGATATTGATATAGATAGATATTGATATAGATAGATATTGATATAGATAGATAGATATAATATGTTTTT
TTTATATATATTACTTTTATTTTCCCTCTTTCTTTCTTTCTTTCTTTTTTTTCCATTCTCTTTTTTTTTAACATTACTACA
ATAATAATTAATATTACTACAATAATAATTAATATTACTACAATAATAATTTATATTACTACAATAATAATTAATATTACT
ACAATAATAATTAATATTACTACAATAATAATTAATATTACTACAATAATAATTTATATTACTACAATAATAATAATAATA
ATGATAATGATAATAATAATAATAATAATAATAATATAATTTAATTACAGCTTAAAGCTTCATCTATCAGTTCATGTATAAATG
AATCAAATAAAATTTTTTTTAAACATAAACCTTGAGGTGGTGCACAGAATTTTACTTTCAATGGCTTTGTAAAATATAAAG
CATCTTGAACTTCCTGTATAGTTAATAATCCAAAACCTACTTGAACTATAGTACCAATTAAAATTCGAATCATATGATATA
AAAAGCGATCTCCTTGAATAACAAATTGATATACATTATTTCTGATTTCCAACATATCAAAAAAATGAATTGTACAAAAAG
GGTTTATTGTTCTTAATTTTTCGGTACCTTTTAATGTTCCTCTAAAACATTCAAAATTGTGATGTCCGATAAAAAGTTTTG
CACATTTTTTTCATTTTTTTCTATATCACATGGCATTATCGAATATGCTGTTTCTCTTTCTTTTTTTTATTATTTGTAATGAAT
AAGATATGTCATTGAAATTTTGTTTTTTCTCTTTATTAAATGGGTTTTTATCTATTTTTAAATTTGTTTTTTCTGTTTCTT
CTTCATTTATATAGTCACTTTCTTCATCCCAATTTTCATTTTTAATCATTTCATTTTTAATCATTTCATTTTTAACAATTT
CATTTTTAACAATTTCATTTTTAACAATTTCATTTTTAACAATTTCATTTTTATCAATTTTCTCTTTATCAATTTCATTTT
TTGTATCTTCATCTTCAATATATTTTTCACCATTATCCTTTTTGGTATAGATTTTCTCAATCGAATTTGATCTATGCATAG
AGATGTCGTTATAGATAGAATTTGACTCATTTTCTTTTCCTTTGTTTTCCTTTTCTTCTTCTTTTTCGTCACTTGTTTCAT
CTTTTTCTTTTTCTAATTCATTTTCATTAAACAGAAAATTTAAGTTTTGTTCATTTAAGATATTCTCTGTTTTTCCATTTT
TCCTTTTACAAAAGATATTACAGTTATAATTGTTTTCACTACTATTTTCACTACATTTTTCATTATTATTTTCACTACTAT
TTTCACTACTATTTTCACTACTATTTACATTATATTTTTCATTACATTTTTCATTACATTTTTCATTACTAGAAATATTACAGTTGTAATTACAAATACTACAATATAATTCCTTATGTTCCCAGTTTTTTTTTTTTTTTAACAAAAGATG
AAAATCGGGGTTCATCGAATAGCTGCCAACAATAATTTCTTTCAGAAGGTTGTCCTGGAATTCGAAGATCTAAATTATATA
AATATATTTTTTCCATAATTATTATATCTTACATTAAAATCATTAAATGGTGCATTTATAATACCTAAAACTTTTATGTCCT
TATTTAATAAACTATTTAAGGCTCTTTTTAATTGACTTAAATTTCCATTACATGGGGTTTTTTTTTCATAAGAAATATATT
GACAAATATATTCTTTTGCATGTACTCCTTTATCAGTTCTACTAACTCCTATAAAATCAAATTGTCTATTTTCAAATTGTC
TATTTTCAAATTGTCTATTTTCAAATTGTCTATTTTCAAATTGTCTATTTTCAAATTGTCTATTTTCAAATTGTCTATTTT
CAAATTGTCTATTTTCAACACGTCTTTCTATATTGTTCTGTTGGTTTTGTTGTTTTTGTTGGTTTAGTTGGTTCTGTTGGT
TTGACCAATCTATTTTCTCTTCATATCCATGTACTTTTAATATACAATTCAAAAGAGTATTTTGTATAGTATTAATTTTTT
TTTCATTATGTAAATTCGTCTTTTCTTCTTTTTCTCTTTTCATTGCATTAAAGTATATATCACTTTTATTTCCTAAACCTG
TCCATCCATTATAATTTGTTCCATCATAATCTACTATTAAAATTATATTTGTATATTTTTTTCTATCATATATTGGTCTCT
CTTTTATTGGATCCTTTATTCCCTTTATTCCATCATATATTCTATTTTTTATTCCATTTATTATTCTATCTTTTATTTTAT
CTTTTATTTTGTCTTTTATTTTATCTTTTATTCCATCTTTTGTTTCTTTCATTTCCATTTTCAATTCTATTTCCTTCTCAC
TTGTACAGTTTATCGTTTCCTCTTCTTTGTAATCATCATTGGTAATTGTTTCATCACTTTCGTTATTTTGTAAAATACTTA
TTTGTGAAGAATCTATATCAAAAGAATCTATTTTAGAAGAATCTATTTTAGAAGAACTTATTTGAAAAACATTTAAAATCT
TATTTGTATTTATATTTGTATTTTTGTTTATATTTGTGTTTATATTTATATTTCTTTTGGTTGAATATGGGATCATTGTTG
GTAATTTTCTGTTTCTTTTGTTTTTCCAATTTTTCAAATTCGGATTTAAAAAAAAGACATTATTTGTCTCAGTAGTTTTAA
TTGTTTTTCTTAAACATTGTGATTTGGCATTTACACTCATATATTTATGAATATAGAAGATAAATATTAATTTTATAGGAA
ATATAAACAAAAGAAAAACAATTTTTTTCATAGTATATTTATTTATGTATGTTTTTTTTTTATTTATCTTTCTACCTTTTT
AAAAAATCCAATACATTTTTTAAAATATTTTTTTAAAATGTTTTTTTTAAAATGTTTTTTTAAAATGTTTTTTTAAAATGT
TTTTTAAAATGTTTTATGGAAATGTTATTTGAAAACACCTTTTAAAAATATTATTTTGATAATTATTTTTTAAAAATACTA
TTTGAAATGCTTTTAAAAATATTATCTGAAATGTTTTTTTATTTTGAAATGTTCTTTAAAAATATGTTTTTAAAAAGAAA
ATATCAGAACAAGGCTAAGATAAGGTCAAGTCAATTCAATTCAGATCAAGTCAAGTCCATTCAACTGGAGTTAATAATTTA
AGTACTTATAAATATAAATAAGTACATCTAAAGTTCTTGTATTTTTTAAAAATAAAAGAATAGTATATAATTTAATAATTT
ATTTAAATATAAAAAAGATTTCTTGAAACATGGTTTTGAACGAAATTTCTGGGAATCTTGTTTTATTACTAAAAGAATTAT
ATTTATAATCTTTGTACGAAACAAAAGAAAAGAAAAGAAAAGAAAAGGGTAAAAACATAAATGTAAATAATTATGTATGTA
TATTTATGTATTTATATATTTTTCTATATATTTCTATTTATATTAAAAATAGCAAAATTCTTCTTTCTTTTATTTTCACAT
TTTCTATTCTTACAACTTTTTTTATTCTCATATAAAGATAATAAACATATAAAAATTATATTCGATTAGATTCTATTAAAA
TAAAACTAAACAAAATATTAGAATCAAAATGAATAAAATCAAATAAAATCGAATAAAATAAAAATATCTTCTCTGTTAATA
ACTAAAATAACACCGATGTGTTTATATAGGTTTTTATTTATATCATTTAATAACTATAGATAAAGAAATATTTTTCTTTTT
TCTTTTTTCATTTCGGAAATAAAATGAGAATAAAAAAGACAACGAAGAAAGCGACAAGTAAAAAAAAAAAAAAAAAAAAA
GAAAAACAAGCGTAGAATATAGTAAATTAAAACAAAACATAATGAACTAAAATATATTAGAGTAACTTAGACATTATTATG
TAAGACCTTATAAAATATTCTTTTGTGAACTAAAAATTTCCAGCAATAACAAAAAAAAATAAATAAAAAATAAAAAGTGAA
TAAAGAGGTTCGGAAATTACCAGTTGGAGATTAAAACGATACTACAATGTGTTTTCAACAAATAATGAATTCAAAAAAACT
ACCATTCGAATTTTTAGACAAACTATTGAAAAAAGGAATATCCAAAAAATATTTATTTTTGTCTCGCTATTCAATGGGTCT
GTATATTTTATATTTACCATCTCATGTTTAACATTTAATAAGAATTTTTTTTTTTTTCAAATCTTACTCTCAATTTGAT
CATAAGTTGTCATATCCACAAAGGTTTATATAAAAGATGATTTTCATTTTTTTGTATTTTTTCATTTTCCCTTTTTTTAAT
TGATTTTATTTACAATTCTTGATTTAAATATATTATATGTTCATATATTTTATAAAAGAGTTTATGTTGTCTTTAATTTTT
TCAATTTTCTTCAATATTGTTTAATATTGTTGAATTTTCAGATTTGCCTTTTTTTTTTTTTTTTTTACTTAATTTATTCTT
ACTTTATTCTTAATATATTCTTAATTTTTAATGGAGAGATAACAAAGTTATGACGTGCATAAATTACAAAAAACGTACATT
ATAAATTGTATTGAAGTAAAAAAAAAAAAAAACAATTGCAAATTATATATTTTCACAACCACACAAACATGAAATTTCGTCC
TATGCTATGTATCACAATATTAAAAGAAAAGAAAAAAAAAAAAAAAAAAATGAAAAGAAATTAATAAAACATGTGAAATGAT
AAATATACATATATAATAAACATATATATGTCGGAGTAATTTAAAAAAAAAGAAAAAAAAAAAAATATATATTGGAGTTT
ATAAAATTCAAAATAGAGAAAAGGAAAAGGAAAATAATACAAATAAATAATATAAACATTAAATAGTAAGAACAGTTCCTT



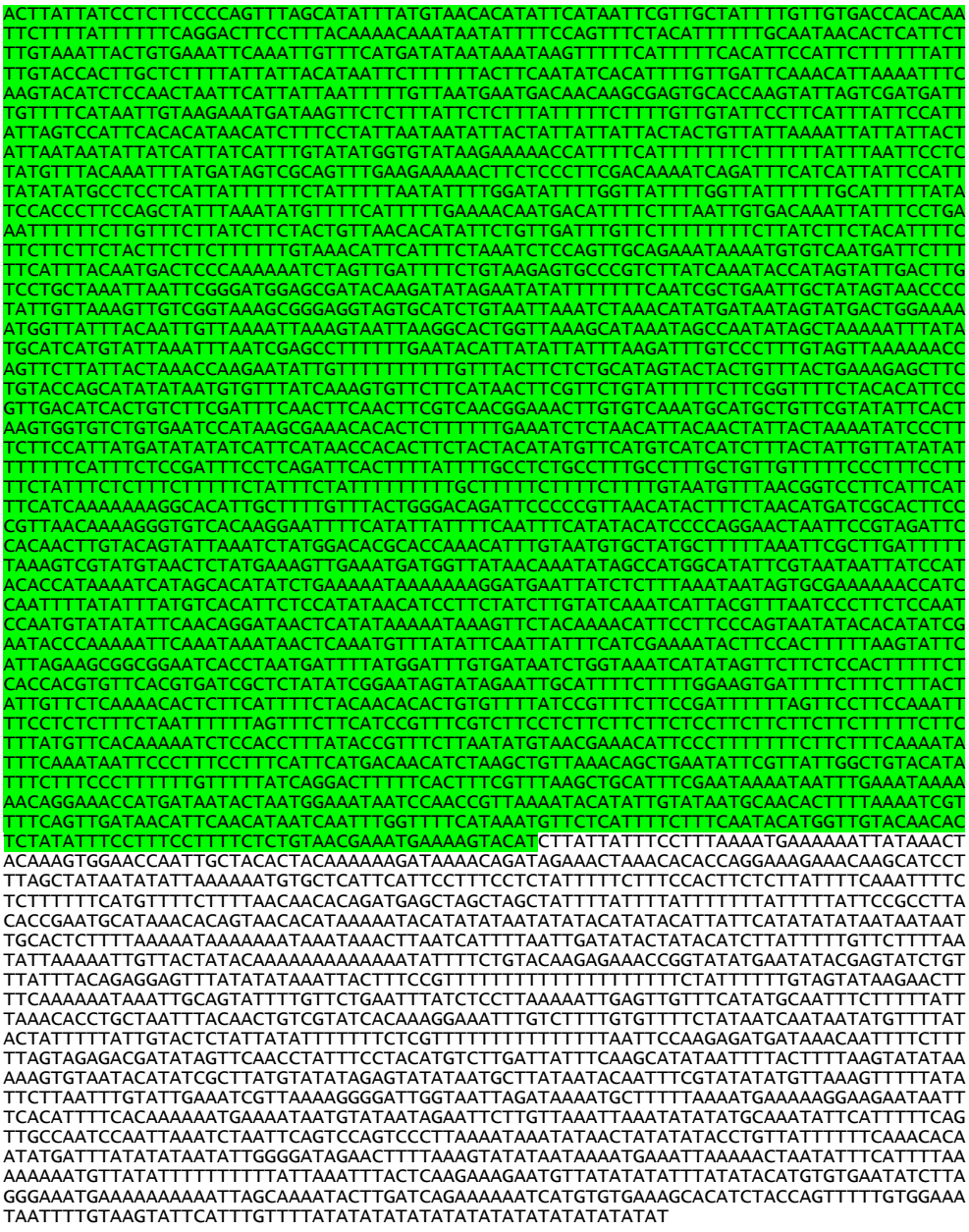


Figure A1. Sequence of contig LSRZ01000339.1. The sequences of ATPase1 (green), 40S ribosomal protein S11 (aqua), and putative tRNA pseudouridine synthase (yellow) genes are highlighted.

>ATPase1, *Haemoproteus tartakovskyi*, genomic DNA
ATGTACTTTTCATTTCGTTACAGAGAAAAGGAAAGGAAATATAGAGTGTTGTACAACCATGTATTGAAAGAAAATGAGAAC
ATTTATGAAAACCAAATTGATTATGTTGAATGTTATCAACTGAAAACGATTTTAAAAGTGTTGCATTATACAATATGTATT
TTAACGGTTGGATTATTTCCATTAGTATTATCATGGTTTCCTGTTTTTTATTTCAAATTATTTTATTCGAAATGCAGCTTA
AACGAAAGTGAAAAAGTCCTGATAAAAACAAAAAAGGGAAAGAAATATGTACAGCCAATAACGAATATTCAGCTGTTTAAC
AGCTTAGATGTTGTCATGAATGAAAGGAAAGGGAATTATTTGAAATATTTTGAAAGAAGAAAAAAAGGGAATGTTTCGTTA
CATATTAAGAAACGGTATAAAGGTGGAGATTTTTGTGAACATAAAGAAGAAAAAGAAGAAGAAGGAGAAGAAGAAGAG
GAAGACGAAACGGATGAAGAAACTAAAAAATTAGAAAGAGAGGAAAATTTGGAAGGAACTAAAAAATCGGAAGAAACGGAT
AAAACACAGTGTGTTGTAGAAAATGAAGAGTGTTTTGAGAACAATAGTAAAGAAAGAAAATCACTTCCAAAAGAAAATGCA
ATTCTATACTATTCCGATATAGAGCGATCACGTGAACACGTGGTGAGAAAAAGTGGAGAAGAACTATATGATTTACCAGAT
TATCACAAATCCATAAAATCATTAGGTGATTCCGCCGCTTCTAATGAATACTTAAAAAGTGGAAGTATTTTCGATGAAATA
ATTGAATATAAACATTTGAGTTATTTATTTGAATTTTTGGGTATTCGATATGTGTATATTACTGGGAAGGAATGTTTTGTA
GAACTTTATTTTATATGAGTTATCCTGTTGAATATATACATTGGATTGGAGAAGGGATTAAACGTAATGATTTGATACAA
GATAGAAGGATGTTATATGGAGAATGTGACATAAATATAAAATTGGATGGTTTTTTCGCACTATTATTTAAAGAGATAATT
CATCCTTTTTTTATTTTTCAGATATGTGCTATGATTTTATGGTGTATGGATAATTATTACGAATATGCCATGGCTATATTT
GTTATAACCATCATTTCAACTTTCATAGAGTTACATACGACTTTAAAAAATCAAGCGAATTTAAAAAGCATAGCACATTAC
AAATGTTTGGTGCGTGTCCATAGATTTAATACTGTACAAGTTGTGGAATCTACGGAATTAGTTCCTGGGGATGTATATGAA
ATTGAAAATAATATGAAAATTCCTTGTGACACCCTTTTGTTAACGGGAAGTGCGATCATGTTAGAAAGTATGTTAACGGGG
GAATCTGTCCCAGTAAACAAAAGCAATGTGCCTTTTTTTGATGAAATGAATGAAGGACCGTTAAACATTACAAAAGAAAAG
AAAAAGCAAAAAAAAATAGAAATAGAAAAAGAAAGAGAAATAGAAAAGGAAAGGGAAAAACAACAGCAAAGGCAAAGGCAG
AGGCAAAATAAAAGTGAATCTGAGGAAATCGGAGAAATGAAAAAAATATATAACAATAGTAAAGATGATGACATGAACATA
TGTAGTAGAAGTGTGGTTATGAATGATATATATCATAATGGAAGAAAGGGATATTTTAGTAATAGTTGTAATGTTAGAGAT

TTCAAAAAAGAGTGTGTTTCGCTTATGGATTCACAGACACCACTTAGTGAATATACGAACAGCATGCATTTGACACAAGTT
TCCGTTGACGAAGTTGAAGTTGAAATCGAAGACAGTGATGTCAACGGAATGTGTAGAAAACCGAAGAAAAATACAGAACGA
AGTTATGAAGAACACTTTGATAAACACATTATATATGCTGGTACAGAAGCTCTTTCAGTAAACAGTAGTACTATGCAGAGA
AGTAAACAAAAAAAAAACAATATTCTTGGTTTAGTAATAAGAACTGGTTTTTTAACTACAAAGGGACAAATCTTAAATAAT
ATAATGTATTCAAAAAAGGCTCGATTAAATTTAATACATGATGCATATAAATTTTTAGCTATATTGGCTATTTATGCTTTA
ACCAGTGCCTTAATTACTTTAATTTTAACAATTGTAAATAACCATTTTTCCAGTCATACTATTATCATATGTTTAGATTTA
ATTACAGATGCACTACCTCCCGCTTTACCGACAACTTTAACAATAGGGGTTACTATAGCAATTCAGCGATTGAAAAAAATA
TATTCTATATCTTGTATCGCTCCATCCCGAATTAATTTAGCAGGACAAGTCAATACTATGGTATTTGATAAGACGGGCACT
CTTACAGAAAATCAACTAGATTTTTTGGGAGTCATTGTAAATGAAAAAGAATCATTGACACATTTTATTTCTGCAACTGGA
GATTTAGAAATGAATGTTTACAAAAAAGAAGAAGTAGAAGAAGAAGAAAATGTAGAAGATAAGAAAAAAAAGAACAAATCA
ACAGAATATGTGTTAACAGTAGAAGATAAGAAACAAGAAAAAATTTCAGGAAATAATTTGTCACAATTAAAGAAAATGTCA
TTGTTTTCAAAAATGAAAACATATTTAAATAGCTGGAAGGGTGGATATAAAAATGCAAAAAATAACCAAAATAACCAAAAT
ATCCAAAATATTAAAAATAGAAAAAATAATGAGGAGGCATATATAAATGGAATAATGATGAAATCTGATTTTGTCGAAGGG
AGAAGTTTTTCTTCAAACTGCGACTATCATAAATTTGTAAACATAGAGGAATTAAATAAAAAAGAAAAAAATGAAAATGGT
TTTTCTTATACACCATATACAAATGATAATGATAATATTATTAATGGTAATAATAATTTTAATAACAGTAGTAATAATAAT
AGTAATATTATTAATAGGAAAGATGTTATGTGTGAATGGACTAATAATGGAATAAATGAAGGAATACAACAAAAGAAAAAT
AAAGAGAATAAAGAGAACTTATCATTTCTTACAATTATGAAAACAAATCATCGACTAATACTTGGTGCACTCGCTTGTTGT
CATTCATTAACAAAAATTAATAATGAATTAGTTGGAGATGTACTTGAAATTTTAATGTTTGAATCAACAAAATGTGATATT
GAAGTAAAAAAGAATTATGTAATAATAAAAGAGCAAGTGGTACAAAATAAAAAAGAATGGAATGTGAAAAATGAAAAACTT
ATTTATTATATCATGAAACAATTTGAATTTCACAGTAATTTACAAAGAATGAGTGTTATTGCAAAAAATGTAGAAACTGGA
AAATATTATTTGTTTTGTAAAGGAAGTCCTGAAAAAATAAAAGAATTGTGTGGTCACAACAAAATAGCAACGAATTATGAA
TATGTGTTACATAAATATGCTAAACTGGGGAAGAGGATAATAAGTATGTCATATAAAGAATTAGATTTAGAAAAAGGAAAT
ATTTTTAATATGGCCCGTAATAAAATGGAAACGGATTTAACATTTCTTGGTTTTTTAATATTTGAAAATAAGTTGAAAACA
TGTGCAAATGATGTTATACATAGTTTGCAGAATTCTGGTTGTCGATGTATTATGTCTACAGGGGATAATATTTTAACTTCT
GTTTCTGTAGCAAAACAATGTGGAATAATTAATGAAAATGTAGATTCGATTATAATTGGAGATATAGTAACAGTCAATAAT
AAAAAGAAAGTTGTGTGGTGTGTTGATCAAAATGACGCTTTATTGAGAGATGGAACAAAATTGAAAGATATAGAAACGAAT
AGCGACCTAGAAATCGGAGCAAAAGAAGTAGAACATGTTGAGAAAAGTGATGGAGAACCTGTTAGTAATGTTTTGAGTTCG
AAGTTATTGTTTAACGACCTGTGTGAACTTTTAATTCATAAAGATCCTAGAAGAACAAGTATTGTGTTATCAGGAGAAGCA
TTTCGTCATTTAAAAAAGCAGTTTGAACAATGTGAATTGCCATATTATGATGACTGTGAGAAAATTACATATGATGCATTA
AAAAATAGTGTTCGAGATAAAGACGATGTGCGAAATCAATATACCTATAAATATACATATAATCATAAATACATAGATACA
GATAATTATAAAAAAGAACAAAACAAATGGGAAACAGTATGTACGAAAAGAAGAATGAACCGAATTACCATTGATAAAGAA
AATAGTGCCGCTTCCTCAAAATGGCATAAAAAAAATATATGTTATAATAAACTTTTGTATAAAATACAGAACAAATTATTT
TCCAATTTCAAATGCTACCCGGAAGAAAAGGAAGGAAGTGAAGAAAAATTAGGTCAAAGAAATGTGTACAGAAGTGAACAG
ATAACATGGAATAAAAATATGAAAGGAGTTAGAGGGGTGAGAGGAGTAAGAGGTATGGCTTCATCTATTTATATGTACCAA
ATGGAAACGAAAAAAGAGAAATGTAGACCTACTGAAAAAAAAATAAAAAAGAAAGGGGAAAGAAAGAGTGTACAGAAGATT
TCGTTAAAAAATGGAAAAAAAAATGAAATGAAAAATGAGCGAATGGGAACCCGATTTAATCTCAAAAGGTTTATAAAAGTT
GCTCAGAAAAGAAAAATTATGGATGTTTTAGTATTTAAAGAACCGACGATTGATGATGAAATAATGAGCGATGGCTTAAAT
GTAAGTTGTTCTATTCGAATGAAAAATTTGAAGCAAAAGAAAATAAACAAATTAAAGAATATAGAAAAAAAAGCTTTTGCA
GGACCGAAACCATTGCTTTCTTTTTGTATAATAAAGAAAACTGATATGTGTCAAATAAAATATCATGGTCCTTCCAATGAC
AAACCTTTAAAAAATTTTATTGTGTATTGGTATAATAATTTTGTTCAATTCAGAGCAAAAAGAAGAAATCAAATAGTACTT
GACAAATCAATGAAGGATCGGAATATGATGCAAGTGAGTTTGCTTGGTAGAAAGTATAGTCGTAGTATGTGCCCATACAAA
AATGATGATGGAGTTAGTGATGATGATGACGATGATGAGGAAGAGGAAGACAAAGAAGAATGTGAATTTGAAACAGAATCG
GAAACCGAATATGCAACAGTAGCAGAAAAAGAAGAAAATCTACAAACAAAAGATGTATATGTAGATATGTATGGAAAAAAG
AGTCGAAGAATTGAAAACAAGAACATGAGGTATAATTTAAATAATCAATCCGAGTGTAATACCGGAAATGTTTTAGAAAAG
AAAAAAACAAGTGATATCATGTATACTACCAAATTGAAAAGAAAGCATCTGAGATGTTCAGAGAAGAAAAAATTCAACAGA
AAAAAAGAACAATCAGAAGAAAGAAAACATCTTTTAAATAGTAACAATAATGGAAACACCCAAATTAATGAGAGTTCTCAA
TTTTTAAAAAATTTTTATGTTCAAAGAATAAAAAGTTACAAAAAACAATTGAAGAATGAAGAACACACAGGGGAAGAATGC
AGTGACTCTTGTTCTCTTTGTACTTCAGTTTCTTCTCTAGGTAGCGATTCCTGTTCATCAAGTTCCTCAACAGAAAATTTA
AAGAAAAGAAGAATAAATGTATATGAAAGGAGGAGAATAAGCCAAATGGGCAGGCGTAGAATGAATCAAATGGGAAGGCGT
GGAATGAATCAAATGGGAAGGCGTGGAATGAATAAAATGGGCAGGAGTGGAAAGAAGAGTGGAACTGAAGAATATATTAGA
AGTGGGTTGATAAACGAATATTTAGGATGCGTAAAACAATCAATTAATCATAAATGTAAAGAAAATACATGTTCGAAATAT
GCACATATACAAAATATAAAAATGTCCATTTATGAATATATTTTAAGAACATGTACTGTGTATGCAAGAATGAATCCATAT
CATAAAACTCAGTTGATATTAAGTTTAAAGAATTTACCTCACAATCCTTTTATTGGGATGTGTGGAGATGGTTCTAATGAT
TGTGGTGCTTTAAAGTCTGCTGATATAGGGATATCATTAAGTAATAATGAATCTTCTATTTGTGCCCCTTTCACTTCTGAC
AACAGTTATTTAGATTCTGTTATAAATGTTTTAATTGAAGGAAGAGCAGCTATTGTAAATTCTTTTCAGTTATTTAAATTT
ATTTCTTTATATTCGATTATACAATGTTCCACTGTGTTAATACTTTATGTTTTGTCAAATAATTTAACAGATAATCAGTTT
ATATATATGGACATATTTACAATTCTTCCTCTTTCTATTTTTATGAGTTGGACATCAGCTAGTAATAAGTTGTCAAATGAT
ATTCCTTTAGGAAAGTTATTTTGCTTTCCTGTTCTTTTCAACTTATATGGACAAATTATGATTCAGCTATTTTTTATTTTC
ATAAGTTTGTTTTTTTTATTTCAGCAACCATTTTATAAAAATGATACTGTGAATAATTACCCTTTAGGCGATACATCCAAC
AAAAAGTTGGTTTGTTCTAAGAATTCTGTCGCCTTTATAGTTTCTTCGTTTCAGGTTCTTTTTATATGTATAGCTCTTAAT
ATAAAAACAAAATGGAGAAAAAGTaagtagaagatttttaacttgatatgtgatattatatgatatgttattatatata
tgatattatatgatattatattatatgatattatattatatgatattatatgttatgatattatatattttgatattatat
gatatgatatatgaggtgataaaatatgagatgttatacatatatatatatttatatattatatatatatggatatatact
taagtttatatctatatgtatatacaaacatatatgtagcctatatatgcattatgttatttatatttcaattgtccatt
tatttttttttttttttttttgcaggcGTTATGACTAATGTTATGTATGTAGGATGGATTTTGATTCTACTTTTTGTTAAT
ACTGCTGTAACTCTTTTATCTAGTGACACCATTTTAATAGGTTTTCTTGTCAAATTTTTGAAAAAATATTTAGGTTTAGTA
AAATTTCCATTATATTTTCGATGGTATTTGCTTATTATTTTATTGTTAAATTTTATTTGTACATTTGTTTTTGAAAAATAT
TATATTCGACATTTAGAGAGAGAAGCAGTGAAAAGGAATTATAAACACACACATATAAATGTACATCCTATTCCTAATAAA
GCAGATACAAGAGTATTTGTCACTTGTTTATGA

>ATPase1, *Haemoproteus tartakovskyi*, protein
MYFSFRYREKERKYRVLYNHVLKENENIYENQIDYVECYQLKTILKVLHYTICILTVGLFPLVLSWFPVFYFKLFYSKCSL
NESEKVLIKTKKGKKYVQPITNIQLFNSLDVVMNERKGNYLKYFERRKKGNVSLHIKKRYKGGDFCEHKEEKEEEEGEEEE
EDETDEETKKLEREENLEGTKKSEETDKTQCVVENEECFENNSKERKSLPKENAILYYSDIERSREHVVRKSGEELYDLPD
YHKSIKSLGDSAASNEYLKSGSIFDEIIEYKHLSYLFEFLGIRYVYITGKECFVELYFYMSYPVEYIHWIGEGIKRNDLIQ
DRRMLYGECDINIKLDGFFALLFKEIIHPFFIFQICAMILWCMDNYYEYAMAIFVITIISTFIELHTTLKNQANLKSIAHY
KCLVRVHRFNTVQVVESTELVPGDVYEIENNMKIPCDTLLLTGSAIMLESMLTGESVPVNKSNVPFFDEMNEGPLNITKEK
KKQKKIEIEKEREIEKEREKQQQRQRQRQNKSESEEIGEMKKIYNNSKDDDMNICSRSVVMNDIYHNGRKGYFSNSCNVRD
FKKECVSLMDSQTPLSEYTNSMHLTQVSVDEVEVEIEDSDVNGMCRKPKKNTERSYEEHFDKHIIYAGTEALSVNSSTMQR
SKQKKNNILGLVIRTGFLTTKGQILNNIMYSKKARLNLIHDAYKFLAILAIYALTSALITLILTIVNNHFSSHTIIICLDL
ITDALPPALPTTLTIGVTIAIQRLKKIYSISCIAPSRINLAGQVNTMVFDKTGTLTENQLDFLGVIVNEKESLTHFISATG
DLEMNVYKKEEVEEEENVEDKKKKNKSTEYVLTVEDKKQEKISGNNLSQLKKMSLFSKMKTYLNSWKGGYKNAKNNQNNQN
IQNIKNRKNNEEAYINGIMMKSDFVEGRSFSSNCDYHKFVNIEELNKKEKNENGFSYTPYTNDNDNIINSNNNFNNSSNNN
SNIINRKDVMCEWTNNGINEGIQQKKNKENKENLSFLTIMKTNHRLILGALACCHSLTKINNELVGDVLEILMFESTKCDI
EVKKNYVIIKEQVVQNKKEWNVKNEKLIYYIMKQFEFHSNLQRMSVIAKNVETGKYYLFCKGSPEKIKELCGHNKIATNYE
YVLHKYAKLGKRIISMSYKELDLEKGNIFNMARNKMETDLTFLGFLIFENKLKTCANDVIHSLQNSGCRCIMSTGDNILTS
VSVAKQCGIINENVDSIIIGDIVTVNNKKKVVWCVDQNDALLRDGTKLKDIETNSDLEIGAKEVEHVEKSDGEPVSNVLSS
KLLFNDLCELLIHKDPRRTSIVLSGEAFRHLKKQFEQCELPYYDDCEKITYDALKNSVRDKDDVRNQYTYKYTYNHKYIDT
DNYKKEQNKWETVCTKRRMNRITIDKENSAASSKWHKKNICYNKLLYKIQNKLFSNFKCYPEEKEGSEEKLGQRNVYRSEQ
ITWNKNMKGVRGVRGVRGMASSIYMYQMETKKEKCRPTEKKIKKKGERKSVQKISLKNGKKNEMKNERMGTRFNLKRFIKV

```
AQKRKIMDVLVFKEPTIDDEIMSDGLNVSCSIRMKNLKQKKINKLKNIEKKAFAGPKPLLSFCIIKKTDMCQIKYHGPSND
KPLKNFIVYWYNNFVQFRAKRRNQIVLDKSMKDRNMMQVSLLGRKYSRSMCPYKNDDGVSDDDDDEEEEDKEECEFETES
ETEYATVAEKEENLQTKDVYVDMYGKKSRRIENKNMRYNLNNQSECNTGNVLEKKKTSDIMYTTKLKRKHLRCSEKKKFNR
KKEQSEERKHLLNSNNNGNTQINESSQFLKNFYVQRIKSYKKQLKNEEHTGEECSDSCSLCTSVSSLGSDSCSSSSSTENL
KKRRINVYERRRISQMGRRRMNQMGRRGMNQMGRRGMNKMGRSGKKSGTEEYIRSGLINEYLGCVKQSINHKCKENTCSKY
AHIQNIKMSIYEYILRTCTVYARMNPYHKTQLILSLKNLPHNPFIGMCGDGSNDCGALKSADIGISLSNNESSICAPFTSD
NSYLDSVINVLIEGRAAIVNSFQLFKFISLYSIIQCSTVLILYVLSNNLTDNQFIYMDIFTILPLSIFMSWTSASNKLSND
IPLGKLFCFPVLFNLYGQIMIQLFFIFISLFFLFQQPFYKNDTVNNYPLGDTSNKKLVCSKNSVAFIVSSFQVLFICIALN
IKTKWRKSKVMTNVMYVGWILILLFVNTAVTLLSSDTILIGFLVKFLKKYLGLVKFPLYFRWYLLIILLLNFICTFVFEKY
YIRHLEREAVKRNYKHTHINVHPIPNKADTRVFVTCL
```

Figure A2. Fasta files of ATPase genomic sequence and amino acid sequence. Intron is in lower case. CDS = join 17290…10705, 10377…10048 of LSRZ01000339.1.

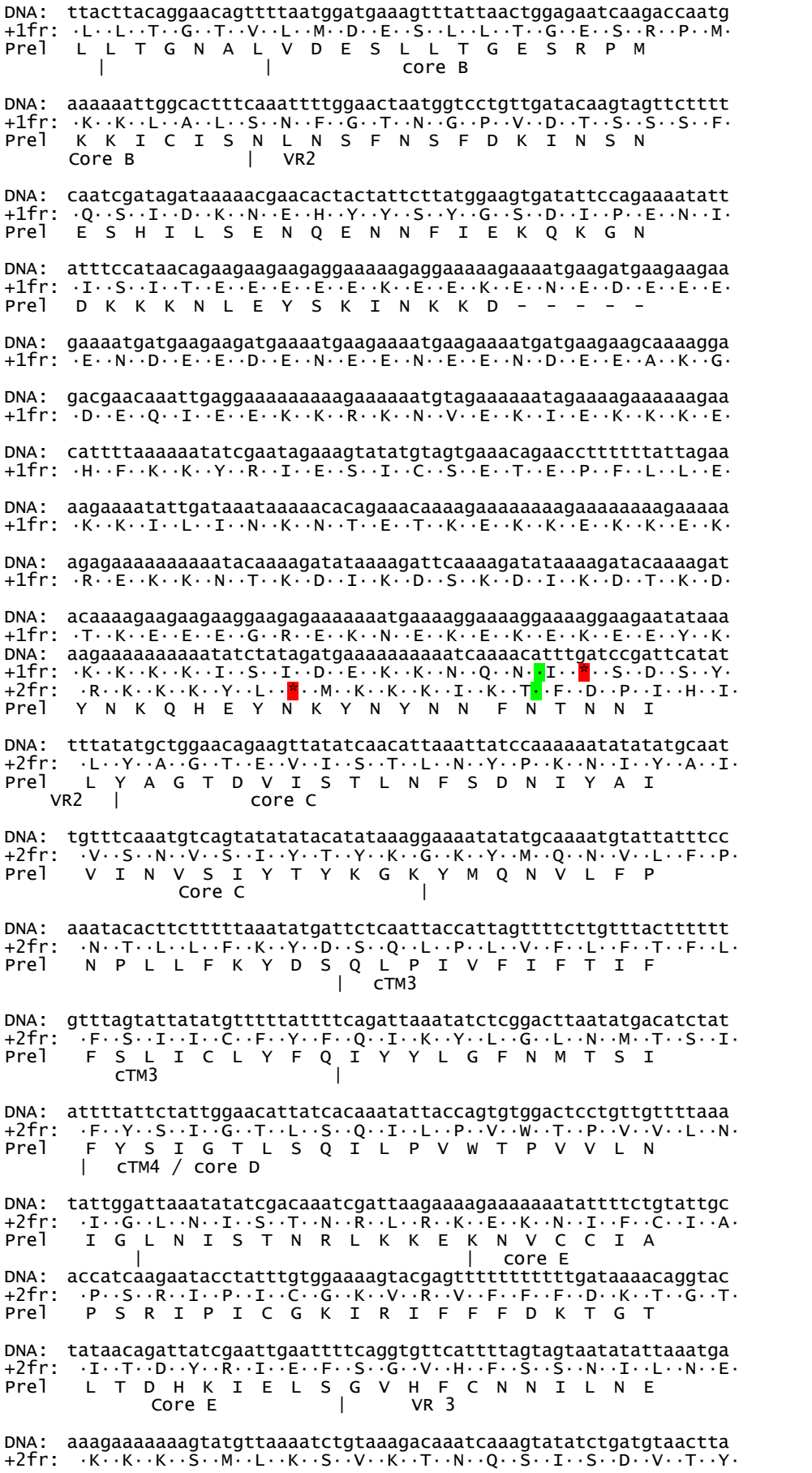
```
>LSRZ01001921.1 Haemoproteus tartakovskyi strain SISKIN1 HtScaffold1921, whole
genome shotgun sequence
ATTATTATTATTATTATTATTATAATTATTATTGTTTTCATATTTTCTATTGAATGAAATTATTTCATTATGTTGTTCATT
TTTTCTTTTATGTATATCTTCATCATTAAAATTTTGAAATGTATTTTCTTTATTAGTTATCCCATTGTTATTTATTATATA
ACTACTATTTCCACTGTTCCATTTTTTTTTTTGAATAATTGTCTACACTTGTTGAAACTGAAATAGGCATAGAAGAAGGTAT
ATATTCATCCATATGTTTGTTCATATTATTTGAATTAAAATATTCTACTATATTATTTGATATAAAATTCGATGGTCCATT
ATTAGTACGATTTTCACTAATATCTTTTATTATATAACTTTCTTTTGAAATATTGTCTTCTATTTCTTCATCACTTATATT
TTGTCTATCATCATATTCTTTTTTATACATTTTTTTTTTTTTTTTCTTATATTTTCAATTTCTTTTTTTTTCTTTTTCTTT
TTCATTTCATCTTCTTCATCTTCTTCATCTTCTTCATCTTCTTCATCTTCTTCATCTTCATCATATTCATCATATTCATCA
TCTTCATCTTCATATTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTCTCGTTCTTCTTTTTTTTGTTGTTGTTCCTGT
TGTTTATCTTTCGAATATACATTTAACAATTTTCTTCGAACCCGAGAAAGATGTTCATATCTAACATCATAAGTTACATCA
GATATACTTTGATTTGTCTTTACAGATTTTAACATACTTTTTTTCTTTTCATTTAATATATTACTACTAAAATGAACACCT
GAAAATTCAATTCGATAATCTGTTATAGTACCTGTTTTATCAAAAAAAAAAAACTCGTACTTTTCCACAAATAGGTATTCTT
GATGGTGCAATACAGAAAATATTTTTTTCTTTTCTTAATCGATTTGTCGATATATTTAATCCAATATTTAAAACAACAGGA
GTCCACACTGGTAATATTTGTGATAATGTTCCAATAGAATAAAATATAGATGTCATATTAAGTCCGAGATATTTAATCTGA
AAATAAAAACATATAATACTAAACAAAAAAGTAAACAAGAAAACTAATGGTAATTGAGAATCATATTTAAAAAGAAGTGTA
TTTGGAAATAATACATTTTGCATATATTTTTCCTTTATATGTATATATACTGACATTTGAAACAATTGCATATATATTTTTT
GGATAATTTAATGTTGATATAACTTCTGTTCCAGCATATAAAATATGAATCGGATCAAATGTTTTGATTTTTTTTTTCATC
TATAGATATTTTTTTTTCTTTTTATATTCTTCCTTTTCCTTTTCCTTTTCATTTTTTTCTCTTCCTTCTTCTTTTTGT
ATCTTTTGTATCTTTTATATCTTTTGAATCTTTTATATCTTTTGTATTTTTTTTTTCTCTTTTTTCTTTTTTTTCTTTTTT
TTCTTTTGTTTCTGTGTTTTTATTTATCAATATTTTCTTTTCTAATAAAAAAGGTTCTGTTTCACTACATATACTTTCTAT
TCGATATTTTTTAAAATGTTCTTTTTTCTTTTCTATTTTTTCTACATTTTTTCTTTTTTTTTCCTCAATTTGTTCGTCTCC
TTTTGCTTCTTCATCATTTTCTTCATTTTCTTCATTTTCATCTTCTTCATCATTTTCTTCTTCTTCATCTTCATTTTCTTT
TTCCTCTTTTTCCTCTTCTTCTTCTGTTATGGAAATAATATTTTCTGGAATATCACTTCCATAAGAATAGTAGTGTTCGTT
TTTATCTATCGATTGAAAAGAACTACTTGTATCAACAGGACCATTAGTTCCAAAATTTGAAAGTGCCAATTTTTTCATTGG
TCTTGATTCTCCAGTTAATAAACTTTCATCCATTAAAACTGTTCCTGTAAGTAATATACAATCACAAGGAATTATCATTTT
TTGATTTATAATTAAAATATCTCCAACAGTTAAATTATTAGAAGGAATTTGTTGAACTATAGAATTTCTATATACAGTAAC
TAAAGTATTATT
```

```
>Reverse Complement of LSRZ01001921.1 Haemoproteus tartakovskyi strain SISKIN1
HtScaffold1921, whole genome shotgun sequence
AATAATACTTTAGTTACTGTATATAGAAATTCTATAGTTCAACAAATTCCTTCTAATAATTTAACTGTTGGAGATATTTTA
ATTATAAATCAAAAAATGATAATTCCTTGTGATTGTATATTACTTACAGGAACAGTTTTAATGGATGAAAGTTTATTAACT
GGAGAATCAAGACCAATGAAAAAATTGGCACTTTCAAATTTTGGAACTAATGGTCCTGTTGATACAAGTAGTTCTTTTCAA
TCGATAGATAAAAACGAACACTACTATTCTTATGGAAGTGATATTCCAGAAAATATTATTTCCATAACAGAAGAAGAAGAG
GAAAAAGAGGAAAAAGAAAATGAAGATGAAGAAGAAGAAAATGATGAAGAAGATGAAAATGAAGAAAATGAAGAAAATGAT
GAAGAAGCAAAAGGAGACGAACAAATTGAGGAAAAAAAAAGAAAAAATGTAGAAAAAATAGAAAAGAAAAAAGAACATTTT
AAAAAATATCGAATAGAAAGTATATGTAGTGAAACAGAACCTTTTTTATTAGAAAAGAAAATATTGATAAATAAAAACACA
GAAACAAAAGAAAAAAAAGAAAAAAAAGAAAAAAGAGAAAAAAAAAATACAAAAGATATAAAAGATTCAAAAGATATAAAA
GATACAAAAGATACAAAAGAAGAAGAAGGAAGAGAAAAAAATGAAAAGGAAAAGGAAAAGGAAGAATATAAAAAGAAAAAA
AAAATATCTATAGATGAAAAAAAAAATCAAAACATTTGATCCGATTCATATTTTATATGCTGGAACAGAAGTTATATCAAC
ATTAAATTATCCAAAAAATATATATGCAATTGTTTCAAATGTCAGTATATATACATATAAAGGAAAATATATGCAAAATGT
ATTATTTCCAAATACACTTCTTTTTAAATATGATTCTCAATTACCATTAGTTTTCTTGTTTACTTTTTTGTTTAGTATTAT
ATGTTTTTATTTTCAGATTAAATATCTCGGACTTAATATGACATCTATATTTTATTCTATTGGAACATTATCACAAATATT
ACCAGTGTGGACTCCTGTTGTTTTAAATATTGGATTAAATATATCGACAAATCGATTAAGAAAAGAAAAAAATATTTTCTG
TATTGCACCATCAAGAATACCTATTTGTGGAAAAGTACGAGTTTTTTTTTTTGATAAAACAGGTACTATAACAGATTATCG
AATTGAATTTTCAGGTGTTCATTTTAGTAGTAATATATTAAATGAAAAGAAAAAAAGTATGTTAAAATCTGTAAAGACAAA
TCAAAGTATATCTGATGTAACTTATGATGTTAGATATGAACATCTTTCTCGGGTTCGAAGAAAATTGTTAAATGTATATTC
GAAAGATAAACAACAGGAACAACAACAAAAAAAAGAAGAACGAGAGAAAGAAGAAGAAGAAGAAGAAGAAGAAGAATA
TGAAGATGAAGATGATGAATATGATGAATATGATGAAGATGAAGAAGATGAAGAAGATGAAGAAGATGAAGAAGATGAAGA
AGATGAAATGAAAAAGAAAAAGAAAAAAAAAAGAAATTGAAAATATAAGAAAAAAAAAAAAAAAAAATGTATAAAAAAGAATAT
GATGATAGACAAAATATAAGTGATGAAGAAATAGAAGACAATATTTCAAAAGAAAGTTATATAATAAAAGATATTAGTGAA
AATCGTACTAATAATGGACCATCGAATTTTATATCAAATAATATAGTAGAATATTTTAATTCAAATAATATGAACAAACAT
ATGGATGAATATATACCTTCTTCTATGCCTATTTCAGTTTCAACAAGTGTAGACAATTATTCAAAAAAAAAATGGAACAGT
GGAAATAGTAGTTATATAATAAATAACAATGGGATAACTAATAAAGAAAATACATTTCAAAATTTTAATGATGAAGATATA
CATAAAAGAAAAAATGAACAACATAATGAAATAATTTCATTCAATAGAAAATATGAAAACAATAATAATTATAATAATAAT
AATAATAATAAT
```

Figure A3. Sequence of contig LSRZ01001921.1 and reverse complement.

```
>ATPase3, Haemoproteus tartakovskyi, partial protein
NNTLVTVYRNSIVQQIPSNNLTVGDILIINQKMIIPCDCILLTGTVLMDESLLTGESRPMKKLALSNFGTNGPVDTSSSFQ
SIDKNEHYYSYGSDIPENIISITEEEEEKEEKENEDEEEENDEEDENEENEENDEEAKGDEQIEEKKRKNVEKIEKKKEHF
KKYRIESICSETEPFLLEKKILINKNTETKEKKEKKEKREKKNTKDIKDSKDIKDTKDTKEEEGREKNEKEKEKEEYKKKK
KISIDEKKNQNFDPIHILYAGTEVISTLNYPKNIYAIVSNVSIYTYKGKYMQNVLFPNTLLFKYDSQLPLVFLFTFLFSII
CFYFQIKYLGLNMTSIFYSIGTLSQILPVWTPVVLNIGLNISTNRLRKEKNIFCIAPSRIPICGKVRVFFFDKTGTITDYR
IEFSGVHFSSNILNEKKKSMLKSVKTNQSISDVTYDVRYEHLSRVRRKLLNVYSKDKQQEQQQKKEEREKEEEEEEEEEY
EDEDDEYDEYDEDEEDEEDEEDEEDEEDEMKKKKKKEIENIRKKKKKCIKKNMMIDKNISDEEIEDNISKESYIIKDISE
NRTNNGPSNFISNNIVEYFNSNNMNKHMDEYIPSSMPISVSTSVDNYSKKKWNSGNSSYIINNNGITNKENTFQNFNDEDI
HKRKNEQHNEIISFNRKYENNNNYNNNNNNN
```

Figure A4. Fasta file of partial protein sequence of ATPase3 from *Haemoproteus tartakovskyi.*


DNA: ttacttacaggaacagttttaatggatgaaagtttattaactggagaatcaagaccaatg
+1fr: ·L··L··T··G··T··V··L··M··D··E··S··L··L··T··G··E··S··R··P··M·
Prel L L T G N A L V D E S L L T G E S R P M
core B
DNA: aaaaaattggcactttcaaattttggaactaatggtcctgttgatacaagtagttctttt
+1fr: ·K··K··L··A··L··S··N··F··G··T··N··G··P··V··D··T··S··S··S··F·
Prel K K I C I S N L N S F N S F D K I N S N
Core B VR2
DNA: caatcgatagataaaaacgaacactactattcttatggaagtgatattccagaaaatatt
+1fr: ·Q··S··I··D··K··N··E··H··Y··Y··S··Y··G··S··D··I··P··E··N··I·
Prel E S H I L S E N Q E N N F I E K Q K G N
DNA: atttccataacagaagaagaagaggaaaaagaggaaaaagaaaatgaagatgaagaagaa
+1fr: ·I··S··I··T··E··E··E··E··E··K··E··E··K··E··N··E··D··E··E··E·
Prel D K K K N L E Y S K I N K K D – – – – –
DNA: gaaaatgatgaagaagatgaaaatgaagaaaatgaagaaaatgatgaagaagcaaaagga
+1fr: ·E··N··D··E··E··D··E··N··E··E··N··E··E··N··D··E··E··A··K··G·
DNA: gacgaacaaattgaggaaaaaaaaagaaaaaatgtagaaaaaatagaaaagaaaaaagaa
+1fr: ·D··E··Q··I··E··E··K··K··R··K··N··V··E··K··I··E··K··K··K··E·
DNA: cattttaaaaaatatcgaatagaaagtatatgtagtgaaacagaaccttttttattagaa
+1fr: ·H··F··K··K··Y··R··I··E··S··I··C··S··E··T··E··P··F··L··L··E·
DNA: aagaaaatattgataaataaaaacacagaaacaaaagaaaaaaaagaaaaaaaagaaaaa
+1fr: ·K··K··I··L··I··N··K··N··T··E··T··K··E··K··K··E··K··K··E··K·
DNA: agagaaaaaaaaaatacaaaagatataaaagattcaaaagatataaaagatacaaaagat
+1fr: ·R··E··K··K··N··T··K··D··I··K··D··S··K··D··I··K··D··T··K··D·
DNA: acaaaagaagaagaaggaagagaaaaaaatgaaaaggaaaaggaaaaggaagaatataaa
+1fr: ·T··K··E··E··E··G··R··E··K··N··E··K··E··K··E··K··E··E··Y··K·
DNA: aagaaaaaaaaaatatctatagatgaaaaaaaaaatcaaaacatttgatccgattcatat
+1fr: ·K··K··K··K··I··S··I··D··E··K··K··N··Q··N··I··*··S··D··S··Y·
+2fr: ·R··K··K··K··Y··L··*··M··K··K··K··I··K··T··F··D··P··I··H··I·
Prel Y N K Q H E Y N K Y N Y N N F N T N N I
DNA: tttatatgctggaacagaagttatatcaacattaaattatccaaaaaatatatatgcaat
+2fr: ·L··Y··A··G··T··E··V··I··S··T··L··N··Y··P··K··N··I··Y··A··I·
Prel L Y A G T D V I S T L N F S D N I Y A I
VR2 core C
DNA: tgtttcaaatgtcagtatatatacatataaaggaaaatatatgcaaaatgtattatttcc
+2fr: ·V··S··N··V··S··I··Y··T··Y··K··G··K··Y··M··Q··N··V··L··F··P·
Prel V I N V S I Y T Y K G K Y M Q N V L F P
Core C
DNA: aaatacacttctttttaaatatgattctcaattaccattagttttcttgtttacttttttt
+2fr: ·N··T··L··L··F··K··Y··D··S··Q··L··P··L··V··F··L··F··T··F··L·
Prel N P L L F K Y D S Q L P I V F I F T I F
cTM3
DNA: gtttagtattatatgtttttatttttcagattaaatatctcggacttaatatgacatctat
+2fr: ·F··S··I··I··C··F··Y··F··Q··I··K··Y··L··G··L··N··M··T··S··I·
Prel F S L I C L Y F Q I Y Y L G F N M T S I
cTM3
DNA: attttattctattggaacattatcacaaatattaccagtgtggactcctgttgttttaaa
+2fr: ·F··Y··S··I··G··T··L··S··Q··I··L··P··V··W··T··P··V··V··L··N·
Prel F Y S I G T L S Q I L P V W T P V V L N
cTM4 / core D
DNA: tattggattaaatatatcgacaaatcgattaagaaaagaaaaaaatattttctgtattgc
+2fr: ·I··G··L··N··I··S··T··N··R··L··R··K··E··K··N··I··F··C··I··A·
Prel I G L N I S T N R L K K E K N V C C I A
core E
DNA: accatcaagaatacctatttgtggaaaagtacgagttttttttttgataaaacaggtac
+2fr: ·P··S··R··I··P··I··C··G··K··V··R··V··F··F··F··D··K··T··G··T·
Prel P S R I P I C G K I R I F F F D K T G T
DNA: tataacagattatcgaattgaattttcaggtgttcattttagtagtaatatattaaatga
+2fr: ·I··T··D··Y··R··I··E··F··S··G··V··H··F··S··S··N··I··L··N··E·
Prel L T D H K I E L S G V H F C N N I L N E
Core E VR 3
DNA: aaagaaaaaaagtatgttaaaatctgtaaagacaaatcaaagtatatctgatgtaactta
+2fr: ·K··K··K··S··M··L··K··S··V··K··T··N··Q··S··I··S··D··V··T··Y·

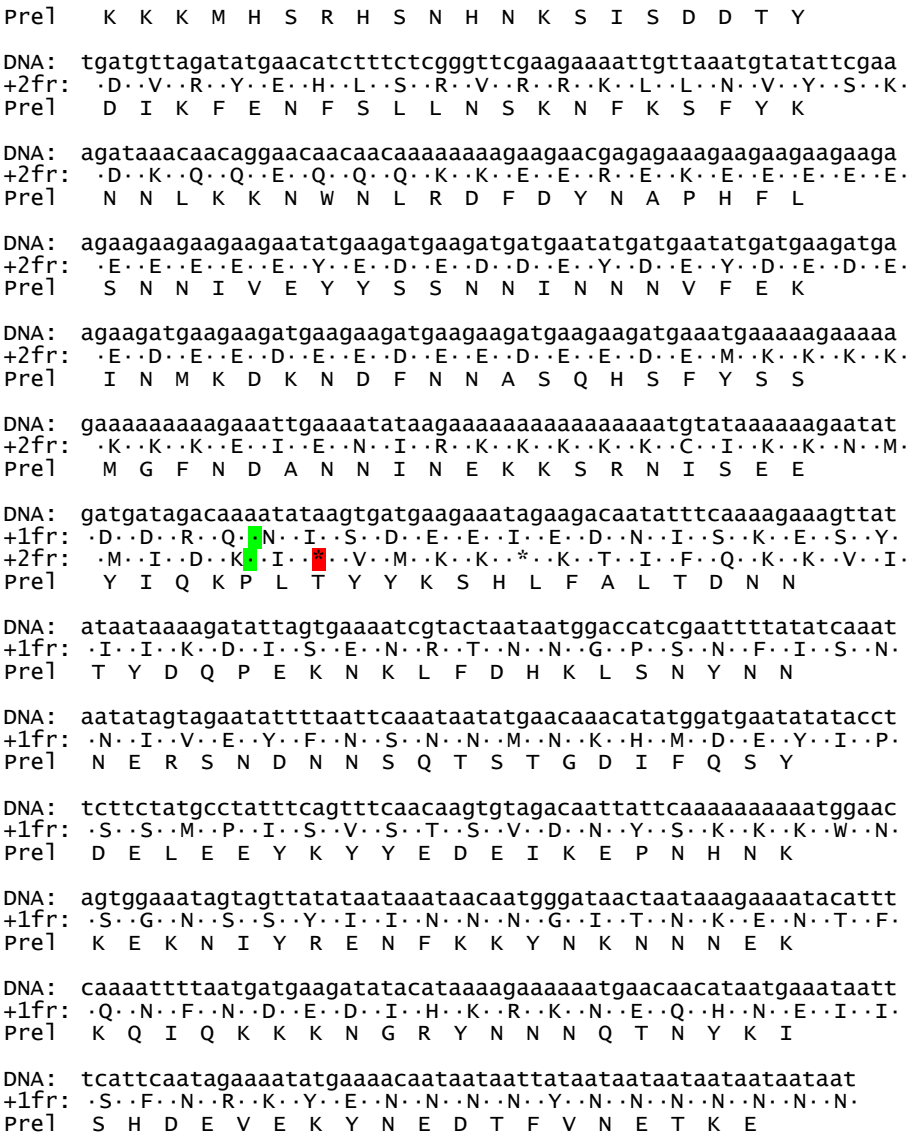


Figure A5. Mapping of possible frame shifts to generate a continuous open reading frame in contig LSRZ01001921.1. Shown is the translation of contig LSRZ01001921.1 with relevant reading frame(s) and an alignment with ATPase3 from *Plasmodium relictum* (Prel). The positions of P-type ATPase cores B–E and variable regions (VR) 2 and 3 are shown below the alignment and demarcated with |. Stop codons that necessitate a frame shift are highlighted in red. The position used to generate the frame shift is highlighted in green.

References

1. Palmgren, M. P-Type ATPases: Many More Enigmas Left to Solve. *J. Biol. Chem.* **2023**, *299*, 105352. https://doi.org/10.1016/j.jbc.2023.105352.
2. Axelsen, K.B.; Palmgren, M.G. Evolution of Substrate Specificities in the P-Type ATPase Superfamily. *J. Mol. Evol.* **1998**, *46*, 84–101. https://doi.org/10.1007/PL00006286.
3. Møller, A.B.; Asp, T.; Holm, P.B.; Palmgren, M.G. Phylogenetic Analysis of P5 P-Type ATPases, a Eukaryotic Lineage of Secretory Pathway Pumps. *Mol. Phylogenet. Evol.* **2008**, *46*, 619–634. https://doi.org/10.1016/j.ympev.2007.10.023.
4. Krishna, S.; Cowan, G.; Meade, J.C.; Wells, R.A.; Stringer, J.R.; Robson, K.J. A Family of Cation ATPase-like Molecules from *Plasmodium Falciparum*. *J. Cell Biol.* **1993**, *120*, 385–398. https://doi.org/10.1083/jcb.120.2.385.
5. Rozmajzl, P.J.; Kimura, M.; Woodrow, C.J.; Krishna, S.; Meade, J.C. Characterization of P-Type ATPase 3 in *Plasmodium Falciparum*. *Mol. Biochem. Parasitol.* **2001**, *116*, 117–126. https://doi.org/10.1016/s0166-6851(01)00319-x.

6. Kimura, M.; Tanabe, K.; Krishna, S.; Tsuboi, T.; Saito-Ito, A.; Otani, S.; Ogura, H. Gametocyte-Dominant Expression of a Novel P-Type ATPase in *Plasmodium Yoelii*. *Mol. Biochem. Parasitol.* **1999**, *104*, 331–336. https://doi.org/10.1016/S0166-6851(99)00153-X.
7. Krishna, S.; Cowan, G.M.; Robson, K.J.; Meade, J.C. *Plasmodium Falciparum*: Further Characterization of Putative Cation ATPases. *Exp. Parasitol.* **1994**, *78*, 113–117. https://doi.org/10.1006/expr.1994.1011.
8. Otto, T.D.; Gilabert, A.; Crellen, T.; Böhme, U.; Arnathau, C.; Sanders, M.; Oyola, S.O.; Okouga, A.P.; Boundenga, L.; Willaume, E.; et al. Genomes of All Known Members of a *Plasmodium* Subgenus Reveal Paths to Virulent Human Malaria. *Nat. Microbiol.* **2018**, *3*, 687–697. https://doi.org/10.1038/s41564-018-0162-2.
9. Hagner, S.C.; Misof, B.; Maier, W.A.; Kampen, H. Bayesian Analysis of New and Old Malaria Parasite DNA Sequence Data Demonstrates the Need for More Phylogenetic Signal to Clarify the Descent of *Plasmodium Falciparum*. *Parasitol. Res.* **2007**, *101*, 493–503. https://doi.org/10.1007/s00436-007-0499-6.
10. Cortés, G.T.; Beltran, M.M.G.; Gómez-Alegría, C.J.; Wiser, M.F. Identification of a Protein Unique to the Genus *Plasmodium* That Contains a WD40 Repeat Domain and Extensive Low-Complexity Sequence. *Parasitol. Res.* **2021**, *120*, 2617–2629. https://doi.org/10.1007/s00436-021-07190-z.
11. DePristo, M.A.; Zilversmit, M.M.; Hartl, D.L. On the Abundance, Amino Acid Composition, and Evolutionary Dynamics of Low-Complexity Regions in Proteins. *Gene* **2006**, *378*, 19–30. https://doi.org/10.1016/j.gene.2006.03.023.
12. Chaudhry, S.R.; Lwin, N.; Phelan, D.; Escalante, A.A.; Battistuzzi, F.U. Comparative Analysis of Low Complexity Regions in Plasmodia. *Sci. Rep.* **2018**, *8*, 335. https://doi.org/10.1038/s41598-017-18695-y.
13. Bajic, M.; Ravishankar, S.; Sheth, M.; Rowe, L.A.; Pacheco, M.A.; Patel, D.S.; Batra, D.; Loparev, V.; Olsen, C.; Escalante, A.A.; et al. The First Complete Genome of the Simian Malaria Parasite *Plasmodium Brasilianum*. *Sci. Rep.* **2022**, *12*, 19802. https://doi.org/10.1038/s41598-022-20706-6.
14. Letunic, I.; Bork, P. Interactive Tree of Life (ITOL) v6: Recent Updates to the Phylogenetic Tree Display and Annotation Tool. *Nucleic Acids Res.* **2024**, *52*, W78–W82. https://doi.org/10.1093/nar/gkae268.
15. Loy, D.E.; Liu, W.; Li, Y.; Learn, G.H.; Plenderleith, L.J.; Sundararaman, S.A.; Sharp, P.M.; Hahn, B.H. Out of Africa: Origins and Evolution of the Human Malaria Parasites *Plasmodium Falciparum* and *Plasmodium Vivax*. *Int. J. Parasitol.* **2017**, *47*, 87–97. https://doi.org/10.1016/j.ijpara.2016.05.008.
16. Votýpka, J.; Modrý, D.; Oborník, M.; Šlapeta, J.; Lukeš, J. Apicomplexa. In *Handbook of the Protists*; Archibald, J.M., Simpson, A.G.B., Slamovits, C.H., Margulis, L., Melkonian, M., Chapman, D.J., Corliss, J.O., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 1–58, ISBN 978-3-319-32669-6.
17. Martinsen, E.S.; Perkins, S.L.; Schall, J.J. A Three-Genome Phylogeny of Malaria Parasites (*Plasmodium* and Closely Related Genera): Evolution of Life-History Traits and Host Switches. *Mol. Phylogenet. Evol.* **2008**, *47*, 261–273. https://doi.org/10.1016/j.ympev.2007.11.012.
18. Burki, F.; Roger, A.J.; Brown, M.W.; Simpson, A.G.B. The New Tree of Eukaryotes. *Trends Ecol. Evol.* **2020**, *35*, 43–55. https://doi.org/10.1016/j.tree.2019.08.008.
19. Wiser, M.F. Protozoa. In *Encyclopedia of Biodiversity (Volume 2)*; Scheiner, S.M.B.T., Ed.; Academic Press: Oxford, UK, 2024; pp. 802–817, ISBN 978-0-323-98434-8.
20. Ciancio, A.; Scippa, S.; Finetti-Sialer, M.; De Candia, A.; Avallone, B.; De Vincentiis, M. Redescription of *Cardiosporidium Cionae* (Van Gaver and Stephan, 1907) (Apicomplexa: Piroplasmida), a Plasmodial Parasite of Ascidian Haemocytes. *Eur. J. Protistol.* **2008**, *44*, 181–196. https://doi.org/10.1016/j.ejop.2007.11.005.
21. Bethke, L.L.; Zilversmit, M.; Nielsen, K.; Daily, J.; Volkman, S.K.; Ndiaye, D.; Lozovsky, E.R.; Hartl, D.L.; Wirth, D.F. Duplication, Gene Conversion, and Genetic Diversity in the Species-Specific Acyl-CoA Synthetase Gene Family of *Plasmodium Falciparum*. *Mol. Biochem. Parasitol.* **2006**, *150B*, 10–24. https://doi.org/10.1016/j.molbiopara.2006.06.004.
22. Tamura, K.; Stecher, G.; Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **2021**, *38*, 3022–3027. https://doi.org/10.1093/molbev/msab120.
23. Cepeda, A.S.; Mello, B.; Pacheco, M.A.; Luo, Z.; Sullivan, S.A.; Carlton, J.M.; Escalante, A.A. The Genome of Plasmodium Gonderi: Insights into the Evolution of Human Malaria Parasites. *Genome Biol. Evol. Evol.* **2024**, *16*, evae027. https://doi.org/10.1093/gbe/evae027.
24. Brayton, K.A.; Lau, A.O.T.; Herndon, D.R.; Hannick, L.; Kappmeyer, L.S.; Berens, S.J.; Bidwell, S.L.; Brown, W.C.; Crabtree, J.; Fadrosh, D.; et al. Genome Sequence of *Babesia Bovis* and Comparative Analysis of Apicomplexan Hemoprotozoa. *PLoS Pathog.* **2007**, *3*, 1401–1413. https://doi.org/10.1371/journal.ppat.0030148.

25. Diesh, C.; Stevens, G.J.; Xie, P.; De Jesus Martinez, T.; Hershberg, E.A.; Leung, A.; Guo, E.; Dider, S.; Zhang, J.; Bridge, C.; et al. JBrowse 2: A Modular Genome Browser with Views of Synteny and Structural Variation. *Genome Biol.* **2023**, *24*, 74. https://doi.org/10.1186/s13059-023-02914-z.
26. Aurrecoechea, C.; Brestelli, J.; Brunk, B.P.; Dommer, J.; Fischer, S.; Gajria, B.; Gao, X.; Gingle, A.; Grant, G.; Harb, O.S.; et al. PlasmoDB: A Functional Genomic Database for Malaria Parasites. *Nucleic Acids Res.* **2009**, *37*, 539–543. https://doi.org/10.1093/nar/gkn814.
27. Tillinghast, J.; Drury, S.; Bowser, D.; Benn, A.; Lee, K.P.K. Structural Mechanisms for Gating and Ion Selectivity of the Human Polyamine Transporter ATP13A2. *Mol. Cell* **2021**, *81*, 4650–4662.e4. https://doi.org/10.1016/j.molcel.2021.10.002.
28. Li, P.; Wang, K.; Salustros, N.; Grønberg, C.; Gourdon, P. Structure and Transport Mechanism of P5B-ATPases. *Nat. Commun.* **2021**, *12*, 3973. https://doi.org/10.1038/s41467-021-24148-y.
29. Schmidt, K.; Wolfe, D.M.; Stiller, B.; Pearce, D.A. Cd2+, Mn2+, Ni2+ and Se2+ Toxicity to *Saccharomyces Cerevisiae* Lacking YPK9p the Orthologue of Human ATP13A2. *Biochem. Biophys. Res. Commun.* **2009**, *383*, 198–202. https://doi.org/10.1016/j.bbrc.2009.03.151.
30. Wiser, M.F. The Digestive Vacuole of the Malaria Parasite: A Specialized Lysosome. *Pathogens* **2024**, *13*, 182. https://doi.org/10.3390/pathogens13030182.
31. Lamarque, M.; Tastet, C.; Poncet, J.; Demettre, E.; Jouin, P.; Vial, H.; Dubremetz, J.-F. Food Vacuole Proteome of the Malarial Parasite *Plasmodium Falciparum*. *Proteomics. Clin. Appl.* **2008**, *2*, 1361–1374. https://doi.org/10.1002/prca.200700112.
32. Phillips, M.A. Polyamines in Protozoan Pathogens. *J. Biol. Chem.* **2018**, *293*, 18746–18756. https://doi.org/10.1074/jbc.TM118.003342.
33. Niemand, J.; Louw, A.I.; Birkholtz, L.; Kirk, K. Polyamine Uptake by the Intraerythrocytic Malaria Parasite, *Plasmodium Falciparum*. *Int. J. Parasitol.* **2012**, *42*, 921–929. https://doi.org/10.1016/j.ijpara.2012.07.005.
34. Bensch, S.; Canbäck, B.; DeBarry, J.D.; Johansson, T.; Hellgren, O.; Kissinger, J.C.; Palinauskas, V.; Videvall, E.; Valkiūnas, G. The Genome of *Haemoproteus Tartakovskyi* and Its Relationship to Human Malaria Parasites. *Genome Biol. Evol.* **2016**, *8*, 1361–1373. https://doi.org/10.1093/gbe/evw081.
35. Böhme, U.; Otto, T.D.; Cotton, J.A.; Steinbiss, S.; Sanders, M.; Oyola, S.O.; Nicot, A.; Gandon, S.; Patra, K.P.; Herd, C.; et al. Complete Avian Malaria Parasite Genomes Reveal Features Associated with Lineage-Specific Evolution in Birds and Mammals. *Genome Res.* **2018**, *28*, 547–560. https://doi.org/10.1101/gr.218123.116.
36. Borner, J.; Pick, C.; Thiede, J.; Kolawole, O.M.; Kingsley, M.T.; Schulze, J.; Cottontail, V.M.; Wellinghausen, N.; Schmidt-Chanasit, J.; Bruchhaus, I.; et al. Phylogeny of Haemosporidian Blood Parasites Revealed by a Multi-Gene Approach. *Mol. Phylogenet. Evol.* **2016**, *94*, 221–231. https://doi.org/10.1016/j.ympev.2015.09.003.
37. Pick, C.; Ebersberger, I.; Spielmann, T.; Bruchhaus, I.; Burmester, T. Phylogenomic Analyses of Malaria Parasites and Evolution of Their Exported Proteins. *BMC Evol. Biol.* **2011**, *11*, 167. https://doi.org/10.1186/1471-2148-11-167.
38. Pacheco, M.A.; Escalante, A.A. Origin and Diversity of Malaria Parasites and Other Haemosporida. *Trends Parasitol.* **2023**, *39*, 501–516. https://doi.org/10.1016/j.pt.2023.04.004.
39. Galen, S.C.; Borner, J.; Martinsen, E.S.; Schaer, J.; Austin, C.C.; West, C.J.; Perkins, S.L. The Polyphyly of *Plasmodium*: Comprehensive Phylogenetic Analyses of the Malaria Parasites (Order Haemosporida) Reveal Widespread Taxonomic Conflict. *R. Soc. Open Sci.* **2018**, *5*, 171780. https://doi.org/10.1098/rsos.171780.
40. Aunin, E.; Böhme, U.; Sanderson, T.; Simons, N.D.; Goldberg, T.L.; Ting, N.; Chapman, C.A.; Newbold, C.I.; Berriman, M.; Reid, A.J. Genomic and Transcriptomic Evidence for Descent from *Plasmodium* and Loss of Blood Schizogony in *Hepatocystis* Parasites from Naturally Infected Red Colobus Monkeys. *PLoS Pathog.* **2020**, *16*, e1008717. https://doi.org/10.1371/journal.ppat.1008717.
41. Tripathi, J.; Stoklasa, M.; Nayak, S.; En Low, K.; Qian Hui Lee, E.; Duong Tien, Q.H.; Rénia, L.; Malleret, B.; Bozdech, Z. The Artemisinin-Induced Dormant Stages of *Plasmodium Falciparum* Exhibit Hallmarks of Cellular Quiescence/Senescence and Drug Resilience. *Nat. Commun.* **2024**, *15*, 7485. https://doi.org/10.1038/s41467-024-51846-0.
42. LaCrue, A.N.; Scheel, M.; Kennedy, K.; Kumar, N.; Kyle, D.E. Effects of Artesunate on Parasite Recrudescence and Dormancy in the Rodent Malaria Model *Plasmodium Vinckei*. *PLoS ONE* **2011**, *6*, e26689.
43. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T.L. NCBI BLAST: A Better Web Interface. *Nucleic Acids Res.* **2008**, *36*, W5–W9. https://doi.org/10.1093/nar/gkn201.

44. Dobson, L.; Reményi, I.; Tusnády, G.E. CCTOP: A Consensus Constrained TOPology Prediction Web Server. *Nucleic Acids Res.* **2015**, *43*, W408–W412. https://doi.org/10.1093/nar/gkv451.
45. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat. Protoc.* **2015**, *10*, 845–858. https://doi.org/10.1038/nprot.2015.053.