A highly rearranged mitochondrial genome in Nycteria parasites (Haemosporidia) from bats

Gregory Karadjian, Alexandre Hassanin, Benjamin Saintpierre, Guy-Crispin Gembu Tungaluna, Frederic Ariey, Francisco J. Ayala, Irene Landau, Linda Duval

A highly rearranged mitochondrial genome in *Nycteria* parasites (Haemosporidia) from bats

Gregory Karadjian [a, b, 1], Alexandre Hassanin[c, 1], Benjamin Saintpierre[d], Guy-Crispin Gembu Tungaluna[e], Frederic Ariey[f], Francisco J. Ayala[g2], Irene Landau[a], Linda Duval[a, 2]

[a] Unité Molécules de Communication et Adaptation des Microorganismes (MCAM, UMR 7245), Sorbonne Universités, Muséum national d'Histoire naturelle, CNRS, CP52, 61 rue Buffon, 75005 Paris, France ; [b] current address, Unité Biologie moléculaire et immunologie parasitaire (UMR 956 ANSES-ENVA-INRA), Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail, 14 rue Pierre et Marie Curie, 94701 Maisons-Alfort Cedex ; [c] Institut de Systématique, Evolution, Biodiversité (ISYEB), UMR 7205 CNRS MNHN UPMC, Muséum National d'Histoire Naturelle, CP51, 57 rue Buffon, 75005 Paris, France ; [d] Unité de Génétique et Génomique des Insectes Vecteurs (CNRS URA3012), Département de Parasites et Insectes Vecteurs, Institut Pasteur, 25-28 Rue du Docteur Roux 75015 Paris, France ; [e] Université de Kisangani, Faculté des Sciences, BP 2012 Kisangani, République Démocratique du Congo ; [f] Laboratoire de Biologie Cellulaire Comparative des Apicomplexes, Faculté de Médicine, Université Paris Descartes - Inserm U1016, CNRS UMR 8104, Cochin Institute, Paris, 75014 France.

Manuscript information: 22 text pages
4 Figures

Word and character counts: 137 words in the abstract
25 035 characters with spaces in the paper

Author contributions: I.L. and L.D. designed research; A.H. and G.C.G.T. collected blood from bats; G.K., I.L. and L.D. performed research; G.K., A.H., B.S., F.A., F.J.A., I.L. and L.D. analyzed data; G.K., A.H., B.S., F.A., F.J.A., I.L. and L.D. wrote the paper.

Author email addresses: gregory.karadjian@anses.fr (G.K.); Hassanin@mnhn.fr (A.H); benjamin.saintpierre@gmail.com (B.S); gembuguycrispin@gmail.com (GCGT); frederic.ariey@yahoo.fr (F.A); fjayala@uci.edu (F.J.A.); ilandau@mnhn.fr (I.L); lduval@mnhn.fr (L.D.);

The authors declare no conflict of interest.

[1] G.K. and A. H. contributed equally to this work.

[2] To whom correspondence may be addressed: fjayala@uci.edu or lduval@mnhn.fr

Data deposition: Relevant genetic sequences were deposited in GenBank and the accession numbers will be provided.

Key words: cytochrome *b*, evolution, phylogeny, *Nycteris*

Abstract

Haemosporidian parasites have mostly and abundantly been described using mitochondrial genes and in particular cytochrome *b*. Failure to amplify the mitochondrial *cytb* gene of *Nycteria* parasites isolated from Nycteridae bats has been recently reported. Bats are hosts to a diverse and profuse array of Haemosporidia parasites which remain largely unstudied. There is a need for obtaining more molecular data from chiropteran parasites. Such data would help to better understand the evolutionary history of Haemosporidia, which notably include the *Plasmodium* parasites, malaria's agents. We use next generation sequencing to obtain the complete mitochondrial genome of *Nycteria* parasites from African *Nycteris grandis* (Nycteridae), and *Rhinolophus alcyone* (Rhinolophidae) and from Asian *Megaderma spasma* (Megadermatidae). We report four new complete mitochondrial genomes, including the two first rearranged mitochondrial genomes within Haemosporidia. Our results open new outlooks regarding potentially undiscovered Haemosporidian diversity.

Significance Statement

Understanding the evolutionary history of *Haemosporidia* parasites would help to understand human malaria evolution. Nevertheless *Haemosporidia* parasite diversity in bats remains largely unstudied. In addition, some cases of unsuccessful PCR amplification of cyt *b*, the most widely used molecular marker to characterize haemosporidian parasites, have been recently reported for *Nycteria* parasites infecting Nycterid bats. Here, we used next generation sequencing to characterize mt genomes of parasites from the *Nycteria* genus. Our results have revealed an unsuspected mitochondrial genome rearrangement within *Haemosporidia*.

Introduction

Haemosporidian blood parasites (phylum Apicomplexa) are vector-borne protozoan parasites that infect multiple vertebrate hosts such as squamates, chelonians, birds and mammals, including humans, in which five species of the genus *Plasmodium* are known to cause malaria in humans (1, 2). The diversity of Haemosporidia has been mostly investigated using mitochondrial genes. In particular, the mitochondrial cytochrome *b* gene (*cytb*) is one of the most widely used genetic markers to characterize the diversity of Haemosporidian parasites and has given rise to many phylogenetic reconstructions to document the evolutionary history of the Haemosporidia (3-5). All Haemosporidia mitochondrial (mt) genomes previously sequenced share the same structure, characterized by a tandemly repeated linear element of approximately 6 kb containing the three protein-coding genes, *cytb*, cytochrome c oxidase subunit 1 (*cox1*) and cytochrome c oxidase subunit 3 (*cox3*) and two highly fragmented (small and large subunits) ribosomal RNA genes (6). In *Plasmodium* species, copy numbers of the mt genome are dozens to hundreds of times more numerous than the nuclear genomes (7, 8). Consequently, mt genes are generally easily amplified using traditional PCR techniques.

Nevertheless, several cases of unsuccessful PCR amplification of mt DNA markers have been recently reported, in particular for *Nycteria* and *Plasmodium* parasites infecting respectively Nycteridae bats and Alaudidae birds (9, 10). Three alternative hypotheses have been proposed to explain these negative results: (1) loss of the mitochondrial genome during the evolution of these parasites; (2) transfer of some mitochondrial genes into the nuclear genome (Numt); (3) technical problems due to use of inefficient or misdesigned primer sets.

Bats are the hosts to a large, diverse and specific array of Haemosporidian parasites (11). Yet, this parasite diversity remains largely unstudied at the molecular and genomic

levels, and this lack of knowledge limits the understanding of the evolutionary history of Haemosporidia, in particular their basal diversification.

Nycteria parasites have been primarily described, based on traditional taxonomy, in African insectivorous bats of two families, Nycteridae and Rhinolophidae (12). The type species of the genus *Nycteria* is *Nycteria medusiformis*, first described from Kenya by Garnham and Heisch in 1953 from the Egyptian slit-faced bat (*Nycteris thebaica capensis*). Later, this species, *N. medusiformis* has been reported in other species of *Nycteris* including the large slit-faced bat, *Nycteris grandis* (13). A total of 6 species have so far been described in the genus *Nycteria* (13).

In order to contribute to the identification and characterization of *Nycteria* parasites in bats and to ascertain the reasons for the failure of PCR amplification in previous studies, we have used deep next generation sequencing (NGS) to assemble the complete mitochondrial genome of four different *Nycteria* species: *N. medusiformis* and *Nycteria* sp. (mixed infection) from *N. grandis* (Nycteridae); *N. gabonensis* from *Rhinolophus alcyone* (Rhinolophidae); and a new species of *Nycteria*, herein described and named *Nycteria heischi* n. sp. from *Megaderma spasma* (Megadermatidae). Our results have revealed an unsuspected mitochondrial genome rearrangement within Haemosporidia.

Results

Three out of 13 *Nycteris grandis* and 4 out of 4 *Rhinolophus alcyone* from a total of 195 bats (6 families) in Democratic Republic of the Congo were found to be infected with *Nycteria* parasites. Four out of 5 *Megaderma spasma* from a total of 90 bats (6 families) in Cambodia were found to be infected with *Nycteria* parasites.

All the three *N. grandis* were co-infected with *Nycteria medusiformis* (14) and an unidentified *Nycteria* sp. The four *R. alcyone* were infected with *Nycteria gabonensis* (13).

Nycteria medusiformis and *Nycteria gabonensis* were identified according to their morphological traits, from *N. grandis* and *R. alcyone* respectively (Fig 1). Thin blood smears of *N. medusiformis* from *N. grandis* (178PBP, 225PB, 226PB MNHN registration numbers) and *N. gabonensis* from *R. sylvestris* (155GG, 376GG, 377GG MNHN registration numbers), collected respectively in 1977 and 1976-77 in Gabon from the Landau collection stored at the Museum national d'Histoire naturelle in Paris, were included in the study for taxonomic examinations. The *Nycteria* parasites from *Megaderma spasma* were characterized and described morphologically and named *Nycteria heischi*, a new species (206YZ MNHN registration number, PXX99 hapantotype collection number).

The morphological characteristics of *Nycteria medusiformis* gametocytes were consistent with the original description (12, 14). Mature gametocytes completely filled the red blood cells (RBC) which were slightly enlarged (8µm instead of 6µm for uninfected RBC, n = 10). In most stages of gametocytes, filaments could be observed at the surface of infected erythrocytes. These filaments became fewer and shorter as the gametocytes matured, and sometimes disappeared at the later stages. Pigment grains were fine and irregularly scattered throughout the cytoplasm. Macrogametocytes were highly chromophilic and appeared dark blue-purple with small and condensed nuclei. Microgametocytes had a pale pink color and the nuclei displayed a small central condensation of chromatin reminiscent of a "*cocarde*" (rosette) (Fig. 1, A-H). The *Nycteria* sp. morphological type, observed in *N. grandis* in a mixed infection, was similar to *N. medusiformis*. Nevertheless, it did not present the highly characteristic filaments observed in *N. medusiformis*. This morphotype differed mostly in size, in the nucleus of the microgametocytes and in the pigment features. However, the number of gametocytes of this *Nycteria* sp. was insufficient for a full description. *N. medusiformis* was highly predominant, while *Nycteria* sp. was less frequent.

With NGS sequencing, we obtained two complete mitochondrial genomes of *Nycteria* from *N. grandis* (273VI), using isegwalker. No other haemosporidian mt genomes were detected in this bat sample by the NGS sequencing. Both mt genomes were 6091 bp-long, with a total of 113661 reads and a mean coverage of 1006. These two *Nycteria* parasites share the same genome organization and their nucleotide sequences differ in only 83 out of 6091 positions (about 1.4 %). Read coverage of 80% and 20% was obtained for the two *Nycteria* mt genomes. Based on the morphological observation frequencies of both *Nycteria* morphotypes, we propose to assign the predominant mt genome to *N. medusiformis* and the less frequent to *Nycteria* sp.

Both *N. medusiformis* and *Nycteria* sp. mt genomes contained the three protein-coding genes, *cytb*, *cox1* and *cox3*, as well as the two highly fragmented ribosomal RNA genes (Fig. 2, B). They displayed a highly rearranged mt genome organization compared to other *Nycteria* parasites and all other Haemosporidian mt genomes sequenced until now. These rearrangements concerned gene order, transcriptional direction (gene inversion at *cox3*), as well as the pattern of high fragmentations of the large subunit (LSU) and the small subunit (SSU) ribosomal RNA (rRNA) genes (Fig. 2, B). However, the rRNA sequences were well conserved between the *Nycteria* parasites from *N. grandis* and all other known Haemosporidian mt genome parasites. The three protein-coding genes, *cox1, cytb* and *cox3* were divergent from that of *N. gabonensis* and *N. heischi* and from all other Haemosporidian parasites.

N. medusiformis mitogenome reconstruction was successfully validated by the amplification of a 1439 bp mt fragment using classical PCR tools with specifically designed primers followed by Sanger sequencing. However, we did not succeeded to amplify and sequence the mt fragment of *Nycteria* sp. using classical PCR and Sanger sequencing. This

classical method favored the amplification and/or sequencing of only *N. medusiformis* parasite in the case of the *N. grandis* mixed infection (15).

Only one species of *Nycteria* was observed in thin blood smears from *Rhinolophus alcyone*. Morphological characteristics were consistent with the original description of *N. gabonensis* gametocytes (13) (Fig. 1, M- P). Briefly, the erythrocyte was deformed and enlarged (mean=9µm, n = 10). It displayed a more or less dark pink color. Mature gametocytes (mean measurements of 7.65 µm, n = 10) were roundish and did not fill the erythrocyte. Microgametocytes had a large well-limited nucleus with a diffuse chromatin and dark pigments, which were mostly seen in the periphery of the nucleus. The macrogametocyte nuclei were smaller and irregular in shape. Only one mitochondrial genome of 5999 bp was reconstructed from the *R. alcyone* (289VI) sample, with a total of 16 361 reads and a mean coverage of 272. This mt genome could be confidently be assigned to *N. gabonensis*.

Nycteria heischi from *Megaderma spasma*

The shape and staining affinities of the *M. spasma* RBCs were not modified by the presence of the *N. heischi* gametocytes. Ring stage invaded normocytes (Fig. 1, I- L). Young gametocytes were ovals and occupied about a quarter of the RBCs. The nucleus was made of two or three chromatin grains of variable size, often located between a bluish relatively large vacuole and a smaller white one. A few very fine grains of pigment could be observed in the cytoplasm. As gametocytes matured, they became rounded. The cytoplasm was light blue or greyish and the nucleus was granular and peripheral, either half-moon shaped in younger stages or larger in older stages. Mature gametocytes of both sexes were pink, of a light color in macrogametocytes and much darker in the microgametocytes. They filled almost completely the RBCs, which were slightly enlarged (mean measurements of about 7 µm, n = 10). There was no “accessory chromatin dot.” The nucleus of the microgametocytes was

arranged as a rosette ("*en cocarde*") comprising a pink more or less homogeneous area devoid of pigment and a central condensed granular zone. The pigment (about 30 grains) was fine, of variable shape and size and dispersed in the cytoplasm. The nucleus of the macrogametocytes was smaller, dense and granular. The cytoplasm was slightly foamy and the pigment was coarser and more abundant than in the microgametocytes.

Nycteria heischi was morphologically related to the group of African *Nycteria* from *Nycteris* bats with the nucleus of the gametocytes "*en cocarde*": *N. medusiformis* (14), *N. houini* and *N. erardi* (13). The schizogonic stages of *N. heischi* are unknown. Gametocytes differ morphologically from (i) *N. medusiformis* by the absence of filaments on the erythrocyte, (ii) *N. houini* by their effect on the RBCs, which stain a dark pink color in the former, but is colorless when seen in *N. heischi*, and by the "rice grains" aspect of the pigment scarcer and irregularly disposed in *N. houini*, and abundant, dispersed and in small grains in *N. heischi*, (iii) *N. erardi* by the "pine needle" pigment scarcer, irregularly disposed and abundant, dispersed and in small grains in *N. heischi*.

Taxonomic summary

Taxon: *Nycteria heischi* n. sp. Karadjian, Landau and Duval, 2016

Type host: *Megaderma spasma* (Linnaeus, 1758)

Type locality: Mondolkiri, Cambodia

Etymology: Named in honor of H. D. Heisch, co-discoverer with P. C. C. Garham of *N. medusiformis*, the type species of the *Nycteria* genus

Hapantotype: one blood film deposited at the collection of the Museum National d'Histoire Naturelle de Paris, N°PXX99 (collection number).

The complete mitochondrial genome of 5989 bp from *Nycteria heischi* (C289) was reconstructed with a total of 64 660 reads and a mean coverage of 734. Both mt-genomes of *N. gabonensis* and *Nycteria heischi*, obtained respectively from *R. alcyone* (289VI) and *M.*

spasma (C289), show the same genome organization as that of all other species of Haemosporidia parasites currently known (Fig. 2, A).

GenBank Accession numbers of the four *Nycteria* mt-genomes are:
Nycteria medusiformis (273VI): KX090645; *Nycteria* sp. (273VI): KX090646, *Nycteria gabonensis* (289VI): KX090647, *Nycteria heischi* (C289): KX090648.

Phylogenetic analyses based on the three datasets (3PCG and *cytb*,) supported the monophyly of the genus *Nycteria* (PP > 0.90) as well as a sister-group relationship between *N. medusiformis* and *Nycteria* sp. from *N. grandis* (PP = 1) (Fig. 3, S1). The branch leading to the two latter species is much longer than other branches of the tree.

The placement of the genus *Nycteria* in deep-level relationships within the Haemosporidia could not be confidently determined.

Discussion

The use of the deep NGS approach has allowed us to discover an unsuspected and highly rearranged mitochondrial genome in *Nycteria* parasites infecting African bats of the family Nycteridae, which had not been successfully amplified and sequenced using classical PCR and Sanger sequencing approaches. Our results reveal that neither the order nor the transcriptional direction of the three protein-coding mitochondrial genes (*cytb*, *cox1* and *cox3*) is conserved within the *Nycteria* genus and the Haemosporidian parasites. The large subunit (LSU) and small subunit (SSU) of the ribosomal RNA (rRNA) gene sequences are well conserved, but their highly fragmented patterns in *N. medusiformis* and *Nycteria* sp. are remarkably different from all other Haemosporidia mt genomes sequenced until now (8). This rearrangement of the mt genomes of *N. medusiformis* and *Nycteria* sp. was discovered by

means of deep NGS sequencing and new primer sets. Outside Haemosporidia order, in other Apicomplexan parasites such as Piroplasma order (*Babesia* and *Theileria* genera), mt-genomes exhibit remarkably diverse structures (6). However, the biological and evolutionary meanings of this mt genome diversity remains to be explored (6).

The use of deep NGS sequencing associated with morphological descriptions might also be an alternative and reliable method to determine the occurrence of co-infections within single hosts (16). Co-infections are often undetectable using standard PCR protocols, resulting frequently in the amplification of a single *cytb* sequence and thus to underestimation of Haemosporidian diversity (15). The PCR underestimation in mixed infection cases could, in turn, lead to unreliable morpho-molecular parasite identifications and phylogenetic analyses, resulting in erroneous interpretations and conclusions. In the present investigation, using NGS, we have confidently identified the occurrence of two distinct *Nycteria* mt genomes present in mixed infection from *N. grandis* (273VI) and two *Nycteria* mono-infections respectively from *R. alcyone* (289VI) and *M. spasma* C289. Based on morphological observations of the two *Nycteria* morphotypes and on the large NGS read coverage of the two mt genomes in *N. grandis*, we associated each *Nycteria* species with its corresponding mitochondrial genome. Thus, the mt genome predominantly found in *N. grandis* could consistently be ascribed to belong to *N. medusiformis* and the minority mt-genome could belong to the *Nycteria* sp. In the same way, we confidently associated *N. gabonensis* and *N. heischi* to their corresponding mt genomes.

Our phylogenetic analyses support the classification of *Nycteria* parasites as a monophyletic group composed of at least two distinct but closely related parasitic lineages, associated with different bat families (12). The distinction between *Nycteria* parasites infecting *Nycteris* and *Rhinolophus* host bats has been proposed previously based on traditional taxonomy (13). Thus, six species have been described and divided into two distinct

groups based on the following morphological sexual-stage characters in blood (gametocytes) and schizont characteristics: (i) *Nycteris* sp. (Nycteridae) parasites with gametocytes arranged as a rosette ("*en cocarde*") (*N. medusiformis*, *N. houini*, and *N. erardi*), (ii) *Rhinolophus* sp. (Rhinolophidae) parasites with gametocytes of the dispersed ("*diffus*") type (*N. gabonensis*, *N. krampitzi*, *N. congolensis*) (13). The *Nycteria heischi* parasites from South East Asian *M. spasma* were morphologically related to the *Nycteria* from African *Nycteris*, with a nucleus of gametocytes "*en cocarde*".

Two phylogenetic placements amongst all major Haemosporidia parasites and evolutionary scenarios based on multi-gene phylogenetic analyses (mostly inferred from *cytb* phylogenetic information) have been previously proposed for the *Nycteria* parasite group. *Nycteria* parasites were placed as a sister clade closely related to the mammalian *Plasmodium*/*Hepatocystis* parasite clade, suggesting that the origin of mammalian Haemosporidia may have been in bats (11). Alternatively, these parasites were placed closer to *Plasmodium* species from lizards, suggesting an alternative scenario such as a host switch event between bat and reptile Haemosporidia parasites in their evolutionary history (9). Such a host-switch event between these terrestrial vertebrate host classes has also been assumed for *Polychromophilus* parasites (infecting Miniopteridae and Vespertilionidae bats) based on close phylogenetic relationships and shared biological traits (schizogony present in the reticulo-endothelial system) (17-20). The mitogenomic phylogenetic analyses presented in the study, using two genetically distantly related Apicomplexan outgroups of Haemosporidia parasites to root phylogenetic reconstructions, did not produced a robust placement of *Nycteria* parasites amongst the Haemosporidia parasites (21, 22). However, the *Nycteria* parasites that have a hepatic schizogony shared this biological trait with mammalian *Plasmodium* and *Hepatocystis* parasites (12, 13).

Next generation sequencing would further open new outlooks regarding unsuspected and unexplored Haemosporidia parasite diversity and would help to understand the systematics and evolution of Haemosporidia parasites.

Materials and Methods

Specimens

A total of 195 bats (6 families) were caught using mist-nets and harp traps during field expeditions in the Orientale province of the Democratic Republic of the Congo (November-December 2013), and 90 bats (6 families) in Cambodia (July 2005 and January 2006). For morphological parasite identification, thin blood smears were used and blood spots were done on Whatman filter paper to be used later for genome analyses.

One sample of each of *N. grandis* (273VI), *R. alcyone* (289VI) and *M. spasma* (C289), infected respectively with *N. medusiformis* and *N*. sp., *N. gabonensis* and *N. heischi*, were used for whole mito-genome sequencing.

Microscopy

Thin blood smears were fixed with absolute methanol, stained with Giemsa solution (10% in phosphate-buffered solution pH = 7.4) for 45 min and then preserved with mounting medium (Eukitt, Polylabo) and a cover slip. Smears were then carefully screened for morphological examinations with a motorized BX63 Upright Olympus Microscope at a magnification of 100x and photographed with an Olympus DP72 Camera (High-Speed 12.8 Megapixel Image Capture).

DNA extraction

Whole genomic DNA from *N. grandis, R. alcyone* and *M. spasma* dried blood spots was extracted using QIAamp DNA Micro Kit (Qiagen) following the manufacturer's instruction handbook. Samples were eluted in 20 µl of Molecular Biology Grade Water. Whole genomic DNAs were quantified using a Qubit ® 2.0 Fluorometer (Life Technologies) and concentrations of 0,997 ng/µl, 0.74 ng/µl and 0.15 ng/µl were respectively obtained for the *N. grandis* sample (including *N. medusiformis* and *N.* sp. parasites), the *R. alcyone* sample (including *N. gabonensis*) and the *M. spasma* sample (including *N. heischi*).

Whole-genome sequencing of Nycteria parasites

Whole mito-genome sequencing was performed on *Nycteria* parasites using the HiSeq 2000 Illumina paired-reads sequencing technology at the Genomics platform at the Pasteur Institute, Paris. The Illumina library preparations were performed using Nextera XT DNA Library Preparation Kits following standard protocols developed by the supplier (Illumina). Briefly, genomic DNA was sheared by nebulization, and sheared fragments were end-repaired and phosphorylated. Blunt-end fragments were A-tailed, and sequencing adapters were ligated to the fragments. Inserts were sized using Agencourt AMPure XP Beads (± 500 bp; Beckman Coulter Genomics) and enriched using 10 cycles of PCR before library quantification and validation. Hybridization of the Nextera DNA libraries to the flow cell and bridge amplification were performed to generate clusters; paired-end reads of 100 cycles were collected on a HiSeq 2000 Illumina. Raw sequence files were filtered using the Fqquality tool, a read-quality filtering software developed by the *gensoft* team at the Pasteur Institute), which enables the trimming of the first and last low-quality bases in reads.

A Perl software named iSeGWalker was newly developed to accomplish a de novo genome reconstruction from the reads file in fastq format provided by Next Generation Sequencing (NGS) data (Illumina). The exact-matching algorithm removed all variations due

to sequencing errors. iSeGWalker was based on a unique and specific seed to Haemosporidian parasites. The sequence “TCTTGAGGCAGTTTGTTCCCTATCTACC”, conserved throughout 72 Haemosporidia species, was chosen as initial seed. The search step is an exact-matching reads selection using a regular expression and a very simple Perl script, reading each sequence one by one. Once all matching reads have been selected, a consensus sequence is determined. Once consensus is determined, a new seed, composed by the 30 last consecutive nucleotides, is obtained and a new search is performed. Reads were then mapped using the new sequence as a reference (BWA software) for all samples in order to identify the parasitic variant present in mixed infection (*N. grandis*).

Assessment of de novo reconstruction

The accuracy of the *de novo* iSeGWalker software was first validated in reconstructing the complete mitochondrial genome of *P. falciparum* 3D7 reference strain from NGS data. PCR and nested-PCR, using newly designed specific primers, were also performed to amplify a mt fragment of 1403 bp from *Nycteria* parasites from *N. grandis* (including rRNA fragments and a portion of *cox1* sequence). PCR was carried out in a total volume of 25 µl consisting of 3 µl of DNA, 5 µl PCR buffer (5x), 1.5 mM MgCl2, 0.2 mM of each dNTPs, 1U of Taq DNA polymerase (Promega) and 0.2mM of each of the primers P5 (5'-AACGCCTGABATGRATGGATAA-3') and P5R (5'- TTCDGGATGWCCAAARAACCAG-3'). PCR cycling conditions were: 2 min at 95°C, 30 sec at 95°C, 30 sec at 53°C and 2 min 30 sec at 72°C for 40 cycles and a final 5 min extension at 72°C. The nested-PCR was performed using 5 µl of the PCR products (diluted 1/10), 5 µl of 5x PCR buffer, 2.5 mM MgCl2, 0.2 mM of each dNTPs, 1U of Taq DNA polymerase (Invitrogen), 0.2mM of the primers P4 (5'-TCCATGTCKTCTCATCGCAG-3') and P6iN (5'- ATAATGTCCATCCAGTTCCACC-3'). Nested-PCR cycling conditions were: 2 min at 95°C, 30 sec at 95°C, 30 sec at 54°C and 1

min 30 sec at 72°C for 40 cycles and a final 5 min extension at 72°C. The PCR products were sent to Eurofins Genomics (France) for sequencing using nested-PCR primers.

Phylogenetic analyses

The four complete mitochondrial genomes of *Nycteria* specifically produced for this study were compared to the mitochondrial sequences available in the nucleotide databases for Haemosporidian parasites. Phylogenetic analyses were performed using two different datasets, named *3PCG* for the protein alignment of three protein-coding genes (1014 amino-acids; 59 taxa, including 50 Haemosporidians and 9 outgroups) and *cytb* for the DNA alignment of the *cytb* gene (1054 nt; 110 Haemosporidians). For protein analyses, the Haemosporidian tree was rooted using nine outgroup species belonging to Eucoccidiorida (*Cyclospora cayetanensis*, *Eimeria acervulina*, *E. maxima*, *E. tenella*, and *Hepatozoon catesbianae*) or Piroplasmida (*Babesia bovis*, *B. microti*, *Cytauxzoon felis*, and *Theileria parva*). GenBank accession numbers of the all genomic and *cytb* sequences are provided in the trees of figures 3 and S1. Complete mitochondrial genomes were aligned using MUSCLE (23) and then further adjusted by eye with Se-Al v2.0a11 (24). For *cytb* datasets, the best-fitting model of sequence evolution (GTR+I+G) was selected under jModelTest 2.1.4 (25) using the Akaike information criterion. Phylogenetic relationships were inferred using Bayesian analyses on MrBayes v3.2.1 (26), and the posterior probabilities (PP) were calculated using four independent Markov chains run for 10,000,000 Metropolis-coupled MCMC generations, with tree sampling every 1000 generations, and a burn-in of 25%. For the protein alignment (*3PCG* dataset), phylogenetic relationships were reconstructed with the same Bayesian approach by applying the mixed model (26).

Acknowledgments

This research study was founded by Labex BCDiv (Biological and Cultural Diversities) and ATM génomique et collections, Muséum national d'Histoire naturelle de Paris. We would like to thank individuals who helped in collecting tissue samples: Raphaël Colombo, Laurent Daudet, Tamas Görföl, Ros Kiri Ing, Jean-François Julien, Prescott Musaba Akawa, and Vuong Tan Tu. We are very grateful to Faustin Toengaho Lokundo, the Rector of the University of Kisangani, for his invitation, and Benjamin Dudu Akaibe, Keunen Hilde, Erick Verheyen, and all other members of the 'Centre de Surveillance de la Biodiversité', for their technical support and administrative assistance.

References

1. Santiago-Alarcon D, Rodríguez-Ferraro A, Parker PG, Ricklefs RE (2014) Different meal, same flavor: cospeciation and host switching of haemosporidian parasites in some non-passerine birds. *Parasit Vectors* 7:286.
2. Liu W et al. (2010) Origin of the human malaria parasite *Plasmodium falciparum* in gorillas. *Nature* 467(7314):420-5.
3. Outlaw DC, Ricklefs RE (2010) Comparative gene evolution in haemosporidian (apicomplexa) parasites of birds and mammals. *Mol Biol Evol* 27(3):537-42.
4. Lacorte GA et al. (2013) Exploring the diversity and distribution of neotropical avian malaria parasites--a molecular survey from Southeast Brazil. *PLoS One* 8(3):e57770.
5. Clark NJ, Clegg SM, Lima MR (2014) A review of global diversity in avian haemosporidians (*Plasmodium* and *Haemoproteus*: Haemosporida): new insights from molecular data. *Int J Parasitol* (5):329-38.
6. Hikosaka K, Kita K, Tanabe K (2013) Diversity of mitochondrial genome structure in the phylum Apicomplexa. *Mol Biochem Parasitol* 188(1):26-33.
7. Vaidya AB, Mather MW (2009) Mitochondrial evolution and functions in malaria parasites. *Annu Rev Microbiol* 63:249-67.
8. Feagin JE et al. (2012) The fragmented mitochondrial ribosomal RNAs of *Plasmodium falciparum*. *PLoS One* (6):e38320.
9. Schaer J et al. (2015) *Nycteria* parasites of Afrotropical insectivorous bats. *Int J Parasitol* 45(6):375-84.
10. Zehtindjiev P et al. (2012) A new morphologically distinct avian malaria parasite that fails detection by established polymerase chain reaction-based protocols for amplification of the cytochrome B gene. *J Parasitol* 98(3):657-65.

11. Schaer J et al. (2013) High diversity of West African bat malaria parasites and a tight link with rodent *Plasmodium* taxa. *Proc Natl Acad Sci USA* 110(43):17415-9.
12. Garnham PCC (1966) Malaria parasites and other Haemosporidia. Blackwell Scientific Publications, Oxford.
13. Rosin G, Landau I, Hugot JP (1978) Considérations sur le genre *Nycteria* (Haemoproteidae) parasite de Microchiroptères africains avec description de quatre nouvelles espèces. *Ann. Parasitol. Hum. Comp* 53(5):447-459.
14. Garnham PCC, Heisch RB (1953) On a new blood parasite of insectivorous bats. *Trans. R.* Soc. *Trop. Med. Hyg* 47(5):357-363.
15. Bernotienė R, Palinauskas V, Iezhova T, Murauskaitė D, Valkiūnas G (2016) Avian haemosporidian parasites (Haemosporida): A comparative analysis of different polymerase chain reaction assays in detection of mixed infections. *Exp Parasitol* 26(163):31-37.
16. Juliano JJ et al. (2010) Exposing malaria in-host diversity and estimating population diversity by capture-recapture using massively parallel pyrosequencing. *Proc Natl Acad Sci USA* 107(46):20138-43.
17. Duval L, Robert V, Csorba G, Hassanin A, Randrianarivelojosia M, Walston J, Nhim T, Goodman SM, Ariey F (2007) Multiple host-switching of Haemosporidia parasites in bats. *Malar J* 6:157.
18. Landau I, Rosin G, Miltgen F, Léger N, Mialhe E (1977) The schizogony of *Polychromophilus* (Haemoproteidae), cosmopolitan parasite of Microchiroptera. *C R Acad Sci Hebd Seances Acad Sci D* 285(15):1311-3.
19. Landau I et al. (1980) Sur le genre *Polychromophilus* (Haemoproteidae, parasite de Microchiropteres). *Annales de Parasitologie*, 55(1): 13-32.

20. Witsenburg F, Salamin N, Christe P (2012) The evolutionary host switches of *Polychromophilus*: a multi-gene phylogeny of the bat malaria genus suggests a second invasion of mammals by a haemosporidian parasite. *Malar J* 11:53.
21. Outlaw DC1, Ricklefs RE (2011) Rerooting the evolutionary tree of malaria parasites. *Proc Natl Acad Sci USA* 108(32):13183-7.
22. Borner J et al. (2015) Phylogeny of haemosporidian blood parasites revealed by a multi-gene approach. *Mol Phylogenet Evol* 94:221-231.
23. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792-7.
24. Rambaw A (2002). http://tree.bio.ed.ac.uk/software/seal/
25. Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 9(8):772.
26. Ronquist F et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61(3):539-42.

Figure Legends

Fig. 1. Microphotographs of gametocytes of *Nycteria* parasites from *Nycteris grandis* (A-H), *Megaderma spasma* (I-L) and *Rhinolophus alycone* (M-P)

A,B: very young gametocytes; C-E: Immature gametocytes; F,G: macro- and microgametocyte; H: old gametocyte of *N. medusiformis* from *Nycteris grandis* (273VI).

I,J: Immature gametocytes; K,L: macro- and microgametocyte of *N. heischi* n. sp. from *Megaderma spasma* (C289).

M,N: Immature gametocytes; O,P: macro- and microgametocyte of *N. gabonensis* from *Rhinolophus alcyone* (289VI).

Fig. 2. The two schematic types of mitochondrial genome organization

A- Typical Haemosporidia mt genome organization, as in *N. gabonensis* and *N. heischi*.

B- *N. medusiformis* and *Nycteria* sp. mt genome organization.

Fig. 3. Mitochondrial genome phylogeny (3PCG) of Haemosporidia parasites including the four *Nycteria* species obtained by MrBayes. The *Nycteria* species group is highlighted in red. Splits with a posterior probability > 0.90 are indicated with asterisk. The tree was rooted with Eucoccidiorida and Piroplasmida.

SI Figure Legend

Fig. S1. *Cytb* phylogeny of the haemosporidian parasites including *Nycteria* sp. from Schaer et al. (2015) and the current study, obtained by MrBayes analysis.

Splits with a posterior probability > 0.80 are indicated. The *Nycteria* group is highlighted in red. The tree is unrooted.

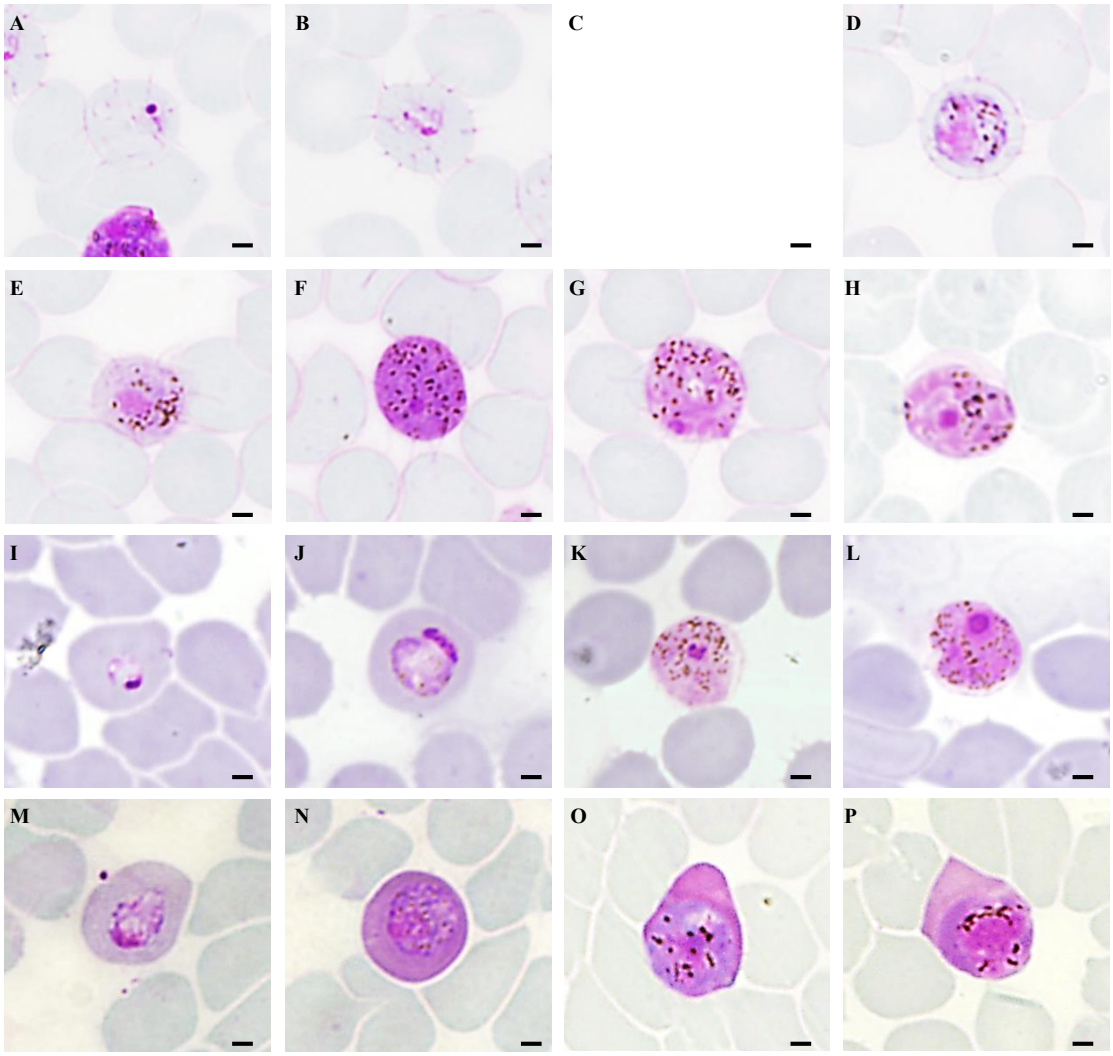

Fig. 1

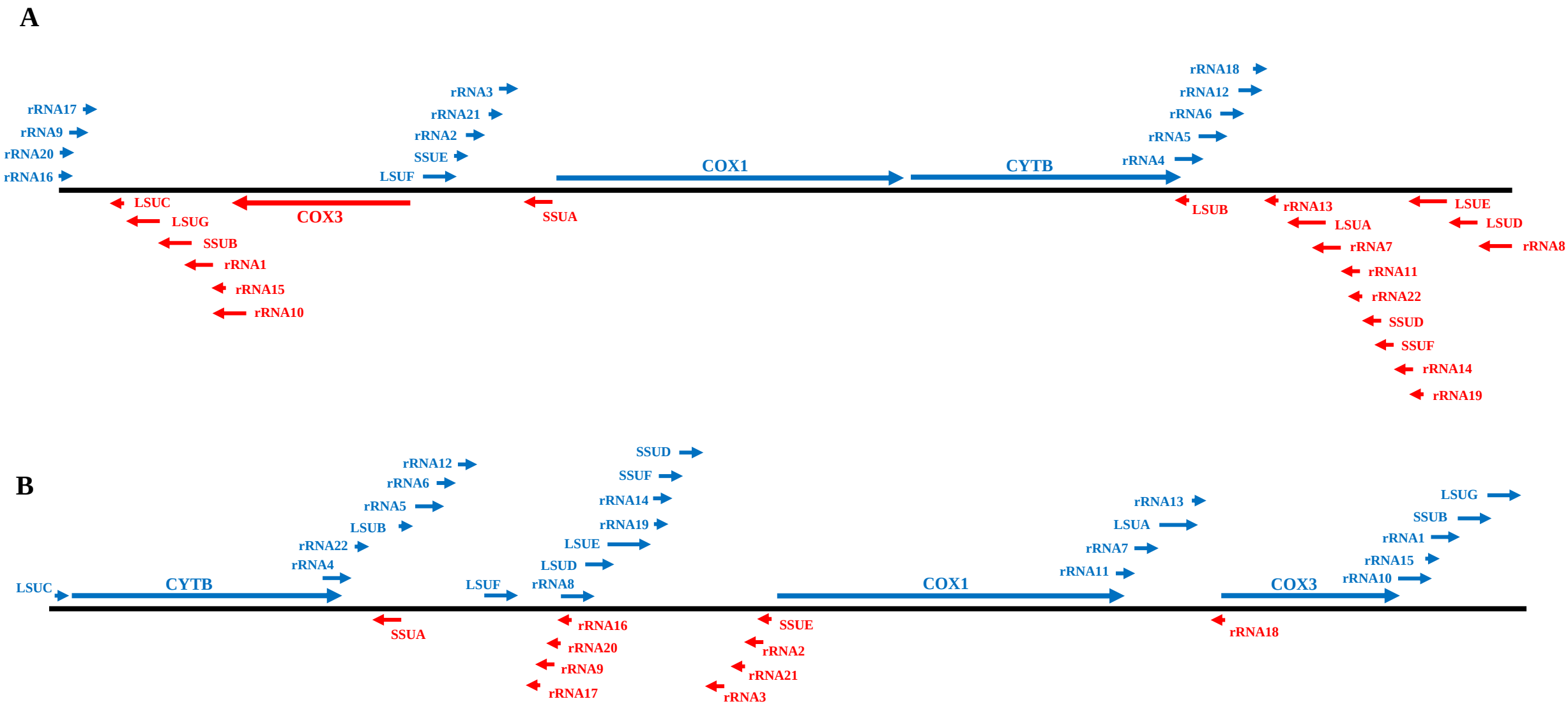
A
rRNA17
rRNA9
rRNA20
rRNA16
rRNA3
rRNA21
rRNA2
SSUE
LSUF
COX1
CYTB
rRNA18
rRNA12
rRNA6
rRNA5
rRNA4
LSUC
LSUG
SSUB
rRNA1
rRNA15
rRNA10
COX3
SSUA
LSUB
rRNA13
LSUA
rRNA7
rRNA11
rRNA22
SSUD
SSUF
rRNA14
rRNA19
LSUE
LSUD
rRNA8
B
SSUD
rRNA12
SSUF
rRNA6
rRNA14
rRNA5
rRNA19
LSUB
LSUE
rRNA22
LSUD
rRNA4
LSUC
CYTB
LSUF
rRNA8
COX1
rRNA13
LSUA
rRNA7
rRNA11
LSUG
SSUB
rRNA1
rRNA15
rRNA10
COX3
SSUA
rRNA16
rRNA20
rRNA9
rRNA17
SSUE
rRNA2
rRNA21
rRNA3
rRNA18

Fig. 2

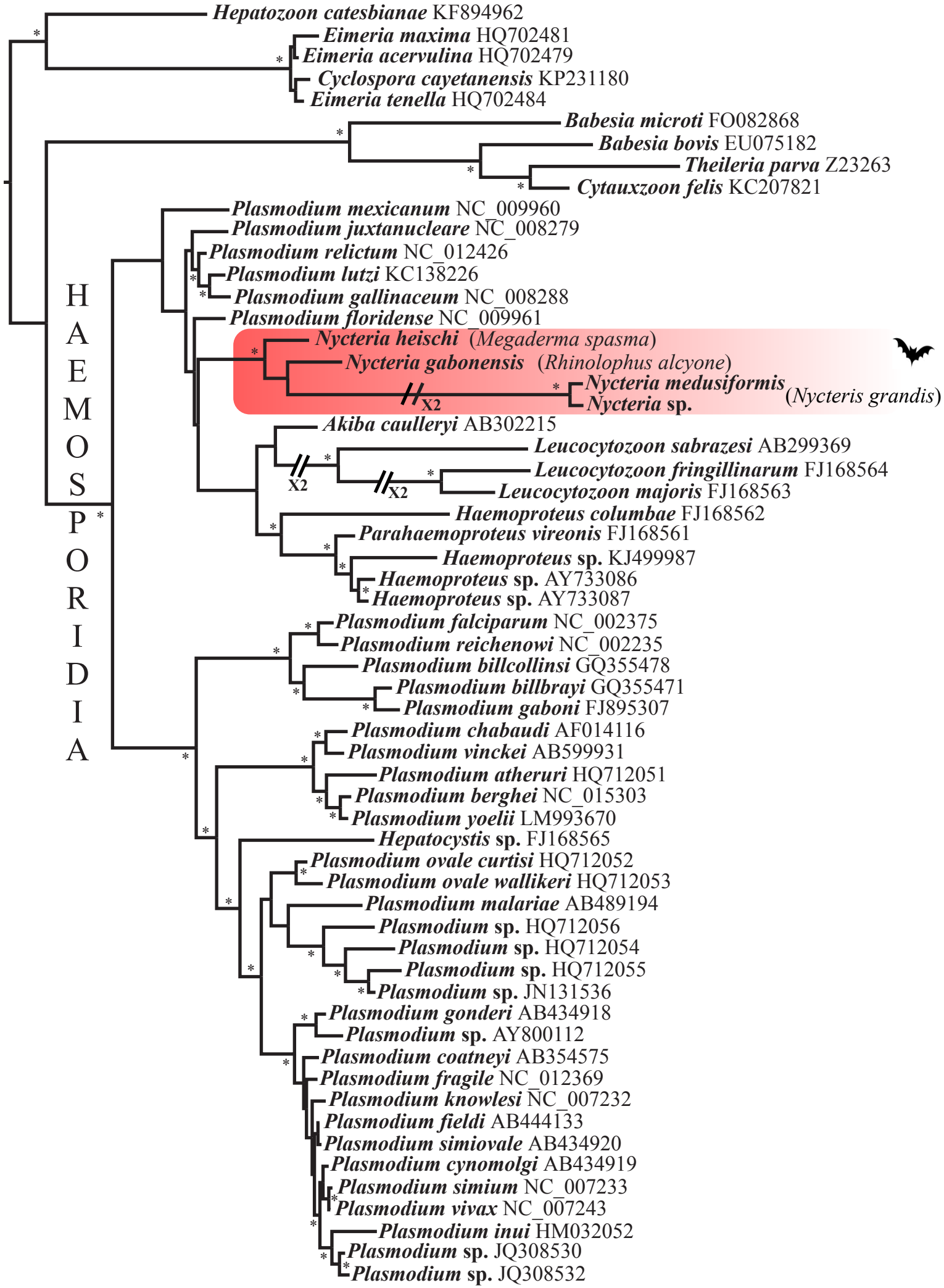
Hepatozoon catesbianae KF894962
Eimeria maxima HQ702481
Eimeria acervulina HQ702479
Cyclospora cayetanensis KP231180
Eimeria tenella HQ702484
Babesia microti FO082868
Babesia bovis EU075182
Theileria parva Z23263
Cytauxzoon felis KC207821
Plasmodium mexicanum NC_009960
Plasmodium juxtanucleare NC_008279
Plasmodium relictum NC_012426
Plasmodium lutzi KC138226
Plasmodium gallinaceum NC_008288
Plasmodium floridense NC_009961
Nycteria heischi (Megaderma spasma)
Nycteria gabonensis (Rhinolophus alcyone)
Nycteria medusiformis
Nycteria sp. (Nycteris grandis)
X2
Akiba caulleryi AB302215
Leucocytozoon sabrazesi AB299369
Leucocytozoon fringillinarum FJ168564
Leucocytozoon majoris FJ168563
Haemoproteus columbae FJ168562
Parahaemoproteus vireonis FJ168561
Haemoproteus sp. KJ499987
Haemoproteus sp. AY733086
Haemoproteus sp. AY733087
Plasmodium falciparum NC_002375
Plasmodium reichenowi NC_002235
Plasmodium billcollinsi GQ355478
Plasmodium billbrayi GQ355471
Plasmodium gaboni FJ895307
Plasmodium chabaudi AF014116
Plasmodium vinckei AB599931
Plasmodium atheruri HQ712051
Plasmodium berghei NC_015303
Plasmodium yoelii LM993670
Hepatocystis sp. FJ168565
Plasmodium ovale curtisi HQ712052
Plasmodium ovale wallikeri HQ712053
Plasmodium malariae AB489194
Plasmodium sp. HQ712056
Plasmodium sp. HQ712054
Plasmodium sp. HQ712055
Plasmodium sp. JN131536
Plasmodium gonderi AB434918
Plasmodium sp. AY800112
Plasmodium coatneyi AB354575
Plasmodium fragile NC_012369
Plasmodium knowlesi NC_007232
Plasmodium fieldi AB444133
Plasmodium simiovale AB434920
Plasmodium cynomolgi AB434919
Plasmodium simium NC_007233
Plasmodium vivax NC_007243
Plasmodium inui HM032052
Plasmodium sp. JQ308530
Plasmodium sp. JQ308532
HAEMOSPORIDIA

Fig. 3

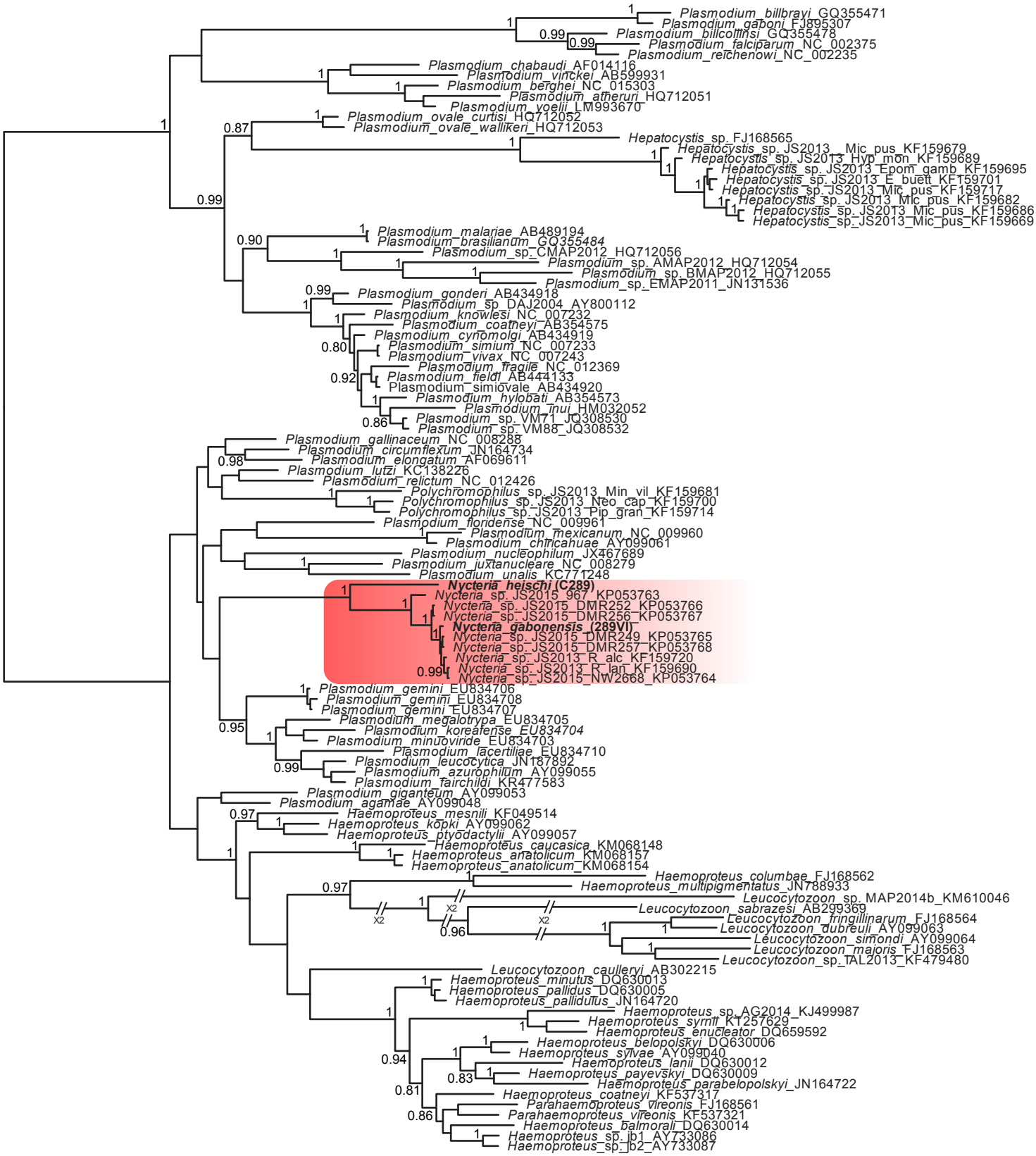
Plasmodium_billbrayi_GQ355471
Plasmodium_gaboni_FJ895307
Plasmodium_billcollinsi_GQ355478
Plasmodium_falciparum_NC_002375
Plasmodium_reichenowi_NC_002235
Plasmodium_chabaudi_AF014116
Plasmodium_vinckei_AB599931
Plasmodium_berghei_NC_015303
Plasmodium_atheruri_HQ712051
Plasmodium_yoelii_LM993670
Plasmodium_ovale_curtisi_HQ712052
Plasmodium_ovale_wallikeri_HQ712053
Hepatocystis_sp._FJ168565
Hepatocystis_sp._JS2013__Mic_pus_KF159679
Hepatocystis_sp._JS2013_Hyp_mon_KF159689
Hepatocystis_sp._JS2013_Epom_gamb_KF159695
Hepatocystis_sp._JS2013_E_buett_KF159701
Hepatocystis_sp._JS2013_Mic_pus_KF159717
Hepatocystis_sp._JS2013_Mic_pus_KF159682
Hepatocystis_sp._JS20T3_Mic_pus_KF159686
Hepatocystis_sp._JS2013_Mic_pus_KF159669
Plasmodium_malariae_AB489194
Plasmodium_brasilianum_GQ355484
Plasmodium_sp._CMAP2012_HQ712056
Plasmodium_sp._AMAP2012_HQ712054
Plasmodium_sp._BMAP2012_HQ712055
Plasmodium_sp._EMAP2011_JN131536
Plasmodium_gonderi_AB434918
Plasmodium_sp._DAJ2004_AY800112
Plasmodium_knowlesi_NC_007232
Plasmodium_coatneyi_AB354575
Plasmodium_cynomolgi_AB434919
Plasmodium_simium_NC_007233
Plasmodium_vivax_NC_007243
Plasmodium_fragile_NC_012369
Plasmodium_fieldi_AB444133
Plasmodium_simiovale_AB434920
Plasmodium_hylobati_AB354573
Plasmodium_inui_HM032052
Plasmodium_sp._VM71_JQ308530
Plasmodium_sp._VM88_JQ308532
Plasmodium_gallinaceum_NC_008288
Plasmodium_circumflexum_JN164734
Plasmodium_elongatum_AF069611
Plasmodium_lutzi_KC138226
Plasmodium_relictum_NC_012426
Polychromophilus_sp._JS2013_Min_vil_KF159681
Polychromophilus_sp._JS2013_Neo_cap_KF159700
Polychromophilus_sp._JS2013_Pip_gran_KF159714
Plasmodium_floridense_NC_009961
Plasmodium_mexicanum_NC_009960
Plasmodium_chiricahuae_AY099061
Plasmodium_nucleophilum_JX467689
Plasmodium_juxtanucleare_NC_008279
Plasmodium_unalis_KC771248
Nycteria_heischi (C289)
Nycteria_sp._JS2015_967_KP053763
Nycteria_sp._JS2015_DMR252_KP053766
Nycteria_sp._JS2015_DMR256_KP053767
Nycteria_gabonensis (289VI)
Nycteria_sp._JS2015_DMR249_KP053765
Nycteria_sp._JS2015_DMR257_KP053768
Nycteria_sp._JS2013_R_alc_KF159720
Nycteria_sp._JS2013_R_lan_KF159690
Nycteria_sp._JS2015_NW2668_KP053764
Plasmodium_gemini_EU834706
Plasmodium_gemini_EU834708
Plasmodium_gemini_EU834707
Plasmodium_megalotrypa_EU834705
Plasmodium_koreafense_EU834704
Plasmodium_minuoviride_EU834703
Plasmodium_lacertiliae_EU834710
Plasmodium_leucocytica_JN187892
Plasmodium_azurophilum_AY099055
Plasmodium_fairchildi_KR477583
Plasmodium_gigantem_AY099053
Plasmodium_agamae_AY099048
Haemoproteus_mesnili_KF049514
Haemoproteus_kopki_AY099062
Haemoproteus_ptyodactylii_AY099057
Haemoproteus_caucasica_KM068148
Haemoproteus_anatolicum_KM068157
Haemoproteus_anatolicum_KM068154
Haemoproteus_columbae_FJ168562
Haemoproteus_multipigmentatus_JN788933
Leucocytozoon_sp._MAP2014b_KM610046
Leucocytozoon_sabrazesi_AB299369
Leucocytozoon_fringillinarum_FJ168564
Leucocytozoon_dubreuli_AY099063
Leucocytozoon_simondi_AY099064
Leucocytozoon_majoris_FJ168563
Leucocytozoon_sp._TAL2013_KF479480
Leucocytozoon_caulleryi_AB302215
Haemoproteus_minutus_DQ630013
Haemoproteus_pallidus_DQ630005
Haemoproteus_pallidulus_JN164720
Haemoproteus_sp._AG2014_KJ499987
Haemoproteus_syrnii_KT257629
Haemoproteus_enucleator_DQ659592
Haemoproteus_belopolskyi_DQ630006
Haemoproteus_sylvae_AY099040
Haemoproteus_lanii_DQ630012
Haemoproteus_payevskyi_DQ630009
Haemoproteus_parabelopolskyi_JN164722
Haemoproteus_coatneyi_KF537317
Parahaemoproteus_vireonis_FJ168561
Parahaemoproteus_vireonis_KF537321
Haemoproteus_balmorali_DQ630014
Haemoproteus_sp._jb1_AY733086
Haemoproteus_sp._jb2_AY733087

Fig. S1