MOLECULAR ECOLOGY RESOURCES WILEY

Contaminations contaminate common databases

Staffan Bensch[1] | **Mizue Inumaru[2]** | **Yukita Sato[2]** | **Larisa Lee Cruz[3,5]** | **Andrew A. Cunningham[4]** | **Simon J. Goodman[5]** | **Iris I. Levin[6]** | **Patricia G. Parker[7]** | **Patricia Casanueva[8]** | **Maria-Angeles Hernández[9]** | **Gregorio Moreno-Rueda[10]** | **Maria-Angeles Rojo[8]**

[1]Department of Biology, MEEL, Lund University, Lund, Sweden

[2]Laboratory of Biomedical Science, Department of Veterinary Medicine, College of Bioresource Sciences, Nihon University, Fujisawa, Japan

[3]Maison de la Télédétection, Montpellier, France

[4]Institute of Zoology, Zoological Society of London, London, UK

[5]School of Biology, Faculty of Biological Sciences, University of Leeds, Leeds, UK

[6]Department of Biology, Kenyon College, Gambier, OH, USA

[7]Department of Biology and Whitney R. Harris World Ecology Center, University of Missouri –St. Louis, St. Louis, MO, USA

[8]Departamento de Ciencias Experimentales, Universidad Europea Miguel de Cervantes, Calle Padre Julio Chevalier, Valladolid, Spain

[9]Departamento de Biología Ambiental, Facultad de Ciencias, Universidad de Navarra, Pamplona, Spain

[10]Departamento de Zoología, Universidad de Granada, Granada, Spain

Correspondence
Staffan Bensch, Lund University, Department of Biology, MEEL, Sweden.
Email: staffan.bensch@biol.lu.se

Funding information
Swedish Research Council, Grant/Award Number: 2017-03937

Abstract

The polymerase chain reaction (PCR) is a very powerful method to detect and identify pathogens. The high sensitivity of the method, however, comes with a cost; any of the millions of artificial DNA copies generated by PCR can serve as a template in a following experiment. If not identified as contaminations, these may result in erroneous conclusions on the occurrence of the pathogen, thereby inflating estimates of host range and geographic distribution. In the present paper, we evaluate whether several published records of avian haemosporidian parasites, in either unusual host species or geographical regions, might stem from PCR contaminations rather than novel biological findings. The detailed descriptions of these cases are shedding light upon the steps in the work process that might lead to PCR contaminations. By increasing the awareness of this problem, it will aid in developing procedures that keep these to a minimum. The examples in the present paper are from haemosporidians of birds, however the problem of contaminations and suggested actions should apply generally to all kinds of PCR-based identifications, not just of parasites and pathogens.

KEYWORDS

Haemoproteus, haemosporidian parasites, PCR contamination, plasmodium, sequence databases

1 | INTRODUCTION

The invention and applications of polymerase chain reaction (PCR) and DNA sequencing have revolutionized research on wild organisms, including parasites and other pathogens. A few of the services provided by these methods include the unambiguous identification of morphologically difficult species, discoveries of cryptic species, delimitation of parasite host range, and resolution of life cycles. Furthermore, the generated DNA sequences can be used for constructing phylogenetic relationships. The importance of these contributions to the ecology and evolution of parasitic organisms cannot be overestimated. To facilitate data sharing it has become standard to report findings of parasites and pathogens based on DNA sequences in open DNA databases such as NCBI, but also in specialist databases delimited on various taxonomic levels or the organisms they infect, such as EuPathDB (Aurrecoechea et al., 2017) and PubMLST (Jolley et al., 2018).

The database MalAvi (http://130.235.244.92/Malavi/index.html) was initiated 10 years ago (Bensch et al., 2009) in order to structure the growing numbers of PCR-based findings of bird blood parasites known as haemosporidians that include malaria parasites. The database presently contains >3,900 unique mitochondrial lineages (based on 478 bp of the cyt b gene) recorded from >1,800 species of birds and 58 species of blood-sucking dipterans. It has become an important resource for analyses of the geographic distribution and host range of these globally distributed parasites (Rivero & Gandon, 2018), with 436 citations (web of science 2020-04-30) and steady annual growth of records. Clearly, the value of a database depends on the quality of the data. Incorrect records may inflate estimates of host species range and geographic distribution of the parasites. In this respect, false positives are more impactful than false negatives. Curated databases such as MalAvi are not resistant to erroneous records. However, as the knowledge of the typical distribution of parasites in host species or in geography are accumulating, unexpected findings can alert us to carefully check and confirm these before publication, and hence keep the database resources as accurate as possible.

In the present paper, we evaluate several unexpected findings of haemosporidian parasites in unusual host species or geographical regions that appear to be the results of PCR contamination or mix up of samples. We would like to emphasize that the authors of these papers are no guiltier of publishing records resulting from contamination than other researchers in the field, as the examples highlighted here probably represent only the tip of the iceberg of a widespread problem. It is only among parasites with restricted and known host and geographic distribution where we have a chance to detect potential cases of contaminations by using this broad approach. For most of the parasites, we do not know where and in what species a finding would be surprising, and hence, any errors among these will pass unnoticed. We believe that the detailed descriptions of these highlighted cases will elucidate what steps might lead to erroneous records and aid in developing procedures that keep these to a minimum. Although the examples in this paper all come from haemosporidians of birds, the problem of contaminations and suggested actions should apply generally to all kinds of PCR-based identifications of parasites and pathogens and, indeed, of other organisms.

2 | CASE STUDIES

In a study with the main focus of describing the species *Haemoproteus iwa* of Pelicaniform birds from the Galapagos (Levin et al., 2011), the authors additionally published sequences of three lineages obtained from blue-footed boobies *Sula nebouxii*, also from the Galapagos. Phylogenetically, these were placed together with lineages of *Parahaemoproteus* of passerine birds, which was somewhat surprising, but as these records were not the focus of the study, the findings were not further discussed in the publication. At the time when the records were added to MalAvi, one of the lineages was shown to have a 100% match to a sequence (EMCIR1) obtained from a cirl bunting *Emberiza cirlus* in Bulgaria (Dimitrov et al., 2010). This was indeed very surprising, as both the locations (12,400 km apart) and the supposed host species (70 million years (Jarvis et al., 2014) were exceptionally distant. It was not immediately clear from the paper which laboratory (University of Missouri – St. Louis, University of Leeds) generated the data from these purportedly infected blue-footed boobies, nor was it obvious at the time that the sequence matched a common parasite that was under study in the same laboratory at the University of Leeds at the same time. Following these first findings, it showed that EMCIR01 was a common parasite of yellow hammers *Emberiza citrinella* in Europe (Dunn et al., 2014) with three closely related (>99% sequence similarity) lineages (EMSPO01, EMRUT01, EMBUC01) in other species of *Emberiza* buntings (Ishtiaq et al., 2007; Nourani et al., 2018; Palinauskas et al., 2013). To explore this remarkable host sharing further, the authors of the Levin et al. (2011) kindly shared samples from four of the infected blue-footed boobies for planned analyses of nuclear gene sequences along with repeated samples of EMCIR1 isolates from European yellow hammers (Huang et al., 2018). However, when the samples were analysed in MEEL, Lund, they showed to be negative with the standard nested protocol for avian haemosporidians (Hellgren et al., 2004) as well as by a protocol amplifying a conserved region of the rRNA of the parasite's mitochondria (Fallon et al., 2003). It seemed to not be a problem of DNA quality as primers for the host mtDNA (Kocher et al., 1989) amplified strong bands. Why were we unable to obtain parasite sequence from these same samples in Lund? The booby sequences had been generated by one of the coauthors of Levin et al. (2011) at the University of Leeds, UK, in the same laboratory where yellow hammer samples infected by EMCIR01 had been processed in parallel (Dunn et al., 2014). Although not proven, it seems more likely that the records of the *Parahaemoproteus* lineages from boobies in Levin et al. (2011) are a result of cross-project contamination rather than true biological findings.

The history of erroneous publications of sequence records from avian haemosporidians may be as old as the research field itself. The

first published cyt b sequence, thought to be from *Haemoproteus columbae* and obtained from a feral pigeon sampled in Venezuela (Escalante et al., 1998), was later shown to be identical to the lineage GRW02 of the morphological species *Plasmodium ashfordi* (Valkiūnas, et al., 2007). The taxonomic mistake is easy to understand since there were no previous cyt b sequences of *Haemoproteus* parasites, and the result can be explained if this pigeon was simultaneously infected by *H. columbae* and *Plasmodium ashfordi* and that the primers (designed based on *Plasmodium* from mammals) preferentially amplified the *Plasmodium* infection. However, the sequence identity to the lineage GRW02 is puzzling. In 1995, Bensch sent a DNA isolate from a great reed warbler *Acrocephalus arundinaceus* infected with what was thought at the time to be an unknown *Haemoproteus* parasite to the authors for them to use in their study of the phylogeny of mammal malaria parasites. Remarkably, this particular sample was later shown to be infected by GRW02 (Bensch et al., 2000). Over the 20 years that have passed since the publication of GRW02 in a pigeon from Venezuela, this lineage now has plenty of records from Europe and Africa, but has never again been recorded in the Americas or in Columbiform birds. We therefore propose that a more plausible explanation is that the record is a result of either PCR contamination or a mix-up of samples in the laboratory procedure.

A lineage-rich cluster of parasites including *Haemoproteus nucleocondensus*, *H. belopolskyi* and *H. payevskyi* are primarily found in warblers of the family Acrocephalidae (Ciloglu et al., 2020; Krizanauskiene et al., 2012; Valkiūnas et al., 2007). The transmission area of the majority of these lineages appear to be restricted to sub-Saharan Africa, as suggested from the observations that only adult individuals of their afrotropical migratory host species carry infections when sampled in Europe. Most of these lineages appear to be strong host specialists (1–3 hosts) with few records in bird species outside this family of songbirds (Bensch et al., 2009). Notable exceptions to this pattern are reports of the lineages MW1, RW1 and GRW01 in dippers *Cinclus cinclus* (Rojo et al., 2015) and MW1 and RW1 in Iberian shrikes *Lanius meridionalis* (Casanueva et al., 2012; Hernández et al., 2017), both studied in Spain. The isolated findings of these lineages in dippers and grey-backed shrikes that are both resident species in Spain and distantly related to Acrocephalidae are surprising in the light of the absence of these lineages in many well sampled and more closely related species in the Iberian peninsula (Mata et al., 2015). Twenty-five of these infected samples were reanalysed in Lund, where primers from other regions of the parasites mtDNA failed amplification. Primers for the hosts amplified strong bands, suggesting that the failure of amplifying the parasites is not explained by degradation of the DNA during the years that passed since the samples were analysed for the original publication. To infer these findings as contamination rather than host shifts and change of transmission area requires a plausible route of how the contamination could have happened. There was a previous study carried out in the same laboratory where the dippers and shrikes were screened, including 149 reed warblers *Acrocephalus scirpaceus*

and 39 sedge warblers *Acrocephalus schoenobaenus*, that reported the parasites MW1 and RW1 (Fernández et al., 2010). Hence, we suggest that the findings of the Acrocephalidae parasites in the dippers and shrikes should be interpreted as due to cross-project contamination.

We have found two more records of these Acrocephalidae-infecting parasites that also might stem from contamination. A study of house sparrows (*Passer domesticus*) from France reported one bird infected by GRW01 (Bonneaud et al., 2006) which is outside the presumed transmission area (Sub-Saharan Africa) of this parasite. Contamination might be a more plausible explanation since the study was carried out in the laboratory in Lund where great reed warblers (the main host of GRW01) were screened for parasites at the same time. Unpublished records of MW1 and the similar lineage (COTCOT03) from quails *Coturnix coturnix* in Spain, deposited in MalAvi, are probably a result of contamination from a parallel screening of *Acrocephalus* warblers which was conducted at the same time (A. D. P. Rodríguez, personal communication).

Plasmodium gallinaceum is a parasite of the domestic chicken *Gallus gallus domesticus* with a documented transmission area restricted to southern Asia (Valkiūnas, 2005), where it can cause substantial mortality in chicken farms. A finding of the cyt b lineage GALLUS01 (the barcoding marker for *P. gallinaceum*) from passerine birds sampled in Japan (Imura et al., 2011) was therefore surprising, but also worrying since this finding indicated that passerines could carry this parasite from winter quarters in southern Asia to Japan, with potential spread to chicken farms. In the same laboratory as the study by Imura et al. (2011), penguins from a zoo in Japan were recently recorded infected with GALLUS01. This was initially interpreted as *P. gallinaceum* having active transmission in Japan, and these unpublished results were discussed at the 4th International Conference on Malaria and Related Haemosporidian Parasites of Wildlife in Beijing, China (Sehgal, 2019). An alternative explanation, however, was suggested: contamination from the PCR positive control that, in this laboratory, routinely was from an isolate of *P. gallinaceum*. This alternative explanation was supported by reanalyses with primers from elsewhere in the parasite mtDNA genome that failed to amplify haemosporidians from the GALLUS01 positive samples in Imura et al. (2011) or from the zoo penguins.

Although most publications do not specify the sample used for positive controls, many studies have probably used *P. gallinaceum* DNA when screening for haemosporidians in wild birds. This is because *P. gallinaceum* was one of the most well studied avian malaria parasites before the molecular revolution, and multiple isolates of DNA from highly infected chickens have been shared and distributed between laboratories worldwide. If available, using *P. gallinaceum* as a positive control was thus a natural first choice in many laboratories when setting up studies of haemosporidians in wild birds. There are a few reports of GALLUS01 from passerine birds or wild-caught mosquitoes sampled outside the known transmission area of *P. gallinaceum* (Kim & Tsuda, 2010; Lacorte et al., 2013; Perkins & Schall, 2002). Whether these are true biological findings or results

of PCR contamination from positive controls remains unknown but would be worth testing for confirmation.

Common lineages are particularly likely to be the source of contamination because every positive PCR is a potential source. Widespread and generalist parasites are therefore highly vulnerable to this problem, particularly because researchers will not be surprised to find them in new host species or geographic regions. Examples of widespread generalists are the lineages SGS1 and GRW04 of the morphospecies *Plasmodium relictum*, recorded in 132 and 82 host species, respectively (MalAvi database, October 2019). However, pinpointing records, if any, as being due to contamination will remain speculative. A report of GRW04 from one single greylag goose *Anser anser* in Germany (Musa et al., 2018) is indicative of contamination as both the host species and the transmission area are surprising. Also, the lineage GRW04 was amplified in a parallel project carried out in the same laboratory (Musa et al., 2019). Whereas the lineage SGS1 is widespread and common in Eurasia and Africa, there are only two published records from the Americas; in a single tree swallow from Canada (Turcotte et al., 2018) and from several species of hummingbirds and passerines at two sites in Peru (Marzal et al., 2015). Since SGS1-infected samples had never been analysed in the laboratory recording the infected tree swallow, contamination seems implausible (D. Garant, personal communication). On the contrary, SGS1 is one of the most abundant lineages amplified in the laboratory in Spain where the Peruvian samples were analysed. However, the authors addressed the contamination concerns by stating in their publication that the "molecular analyses were repeated in three independent laboratories to verify the validity and reproducibility of detection of the SGS1 lineage" (Marzal et al., 2015).

3 | RECOMMENDATIONS AND CONCLUSION

Completely avoiding PCR contamination in publications would require substantial resources and efforts, e.g., by confirming each infection with multiple primer sets, from independent extractions from the same sample or by analyses in different laboratories. However, we can substantially reduce the reporting in publications of erroneous results due to PCR contamination by being open to the possibility that these may happen, even in the best organized laboratories, and reanalyse any samples that produce unexpected or otherwise suspicious results. In Table 1, we provide a checklist of questions to help evaluate whether a finding is the result of contamination; the risk is increasing with the number of these questions that can be answered with a "yes".

The first measure to take is to reduce the risk of contamination and to monitor background levels of contamination in the laboratory. The detection power of nested PCR comes with a cost; the PCR products from the first PCR can easily contaminate samples in the second PCR. It is therefore strongly advisable to allocate separate locations (preferably separate rooms) and pipettes for DNA extraction, PCR setup and handling of PCR-products. For monitoring the contamination in the laboratory, the use of negative controls (NTC) should never be compromised. It is important to emphasize that using only one NTC per experiment will mainly test if the reagents in the master mix (including the water) are contaminated. To monitor the presence of low levels of background contamination (e.g., aerosols of previous PCR products that might precipitate in any of the tubes), it is advisable to use multiple NTCs (e.g., 1 out of 8) which will increase the chance of detecting background contamination. If NTCs repeatedly are found positive, determining the lineage of these by sequencing will provide knowledge on which contaminants are present and could possibly help with detecting the source of the contamination.

The way the NTCs are organized in the experiment might also affect the chance of detecting low-level background contamination. This is illustrated by a study carried out in the laboratory in Lund that investigated 382 samples from birds from the Azores that had overall low rates of infection (Hellgren et al., 2011). All of the NTCs (one out of eight samples) were negative as were most of the samples from 10 species of birds. Five samples from four species, however, were positive for the lineage PARUS1, a common parasite of tits in Europe, but not previously recorded in the species being tested. These samples proved to be blood-smear negative and, since

Questions	Alert level
Does the host species belong to a different family than previously known host species?	Some
Have microscopic analyses of blood smears failed to find infected cells?	Some
Has the lineage been encountered previously in recent projects in the laboratory?	Medium
Is the main host species of the lineage included in the same experiment?	Medium
If generated by nested PCR, were the setup for the second PCR done at the same workspaces and/or pipettes as the first PCR?	High
Were the PCRs done by beginners in the laboratory, e.g., students?	High
Was the sample run next to another sample shown to be positive for the lineage?	High
Is the lineage the same as in the sample of the positive control?	High
Does the sample fail to amplify with general screening primers?	High

TABLE 1 Checklist alerting whether findings of haemosporidian cytochrome b lineages in novel host species may require additional examination before publications to verify that these are not the result of PCR-contamination. The more of the questions that can be answered by "yes" the more likely that the finding is a result of contamination

the analyses of re-extracted DNA failed amplification, the findings were excluded in the final publication as they were suspected to be due to contamination from a previous project conducted in the same laboratory on blue tits (*Cyanistes caeruleus*) infected with PARUS1 (Stjernman et al., 2008). The contaminated samples were all on the same row of the 96-plate (row D) whereas the 12 NTC were placed on the bottom (row H). Hence, it was thought that the contamination was from one of the channels in the multipipette (position 4 from the left), which could explain why it did not show up in the NTCs. Arranging the NTCs cross-wise on the plate (A-H) would have improved the chances of picking up such a source of contamination.

Many laboratories nowadays use separate protocols for screening (Ciloglu et al., 2019; Fallon et al., 2003) and for the sequencing (Hellgren et al., 2004). Disagreement between repeated PCRs (with the same or different primers) is however expected if infection intensities are low (around the detection limit of the PCR), so it may not be a sign of contamination (Bensch & Hellgren, 2020). However, cases where sequences are obtained from samples that are repeatedly negative by the screening primers should be carefully evaluated. If the lineage is new to the host species (and of a lineage that previously has been amplified in the laboratory), it should be verified by additional investigation. The most convincing test would be to have a subsample of the original sample analysed in a different laboratory; however, that might be difficult to arrange. Because contamination is likely to be from a PCR product of the primers used for sequencing, designing new primers outside this region is a viable alternative. Whole mtDNA genome sequences are available from an increasing number of haemosporidians (Pacheco et al., 2018). Hence, designing primers for novel regions of the mtDNA genome is nowadays quite straightforward.

Samples selected to be used for positive controls are typically from highly infected isolates as one wants a robust sample when checking the success of the assay. However, the repeated use of highly infected samples increases the risk of generating PCR contaminants that may carry over to other samples in the same experiment or form background contamination in the laboratory which could affect later projects. The ignorance of positive controls being the potential source of contamination dominates studies of avian haemosporidians; <5% of the papers included in the MalAvi database have reported the lineage(s) of the sample(s) used as positive controls. Insightful selection of samples for positive controls may also highlight cases of potential contamination. For example, it would be more informative to use samples from a known host-specific parasite when screening samples from other bird species than using a sample from known generalist parasites. The finding of the former in an unexpected host species would call for checking the sample again, using a new DNA extract or primers from a different region of the parasite's genome. For this reason, *P. gallinaceum* would be an excellent positive control in studies of nongalliform birds.

Collecting blood smears in parallel with blood for molecular analyses is strongly recommended as it allows for confirming the presence of a parasite by microscopic analyses, which effectively removes concerns of contamination. However, PCR-positive but blood smear negative samples are expected to be relatively frequent due to submicroscopic infection intensities or aborted development of the parasite (Moens et al., 2016). The latter refers to situations where the parasite can infect and replicate in internal organs, but fails to infect blood cells and hence cannot complete transmission as this requires formation of gametocytes in the blood. Therefore, microscopy can verify the presence of a parasite in a sample but if negative, cannot separate PCR contamination from low level or aborted infections.

Finally, mixing up samples in the field or in the laboratory can lead to wrongly assigned parasites to host species. If a tube was mislabelled in the field, analyses of re-extracted DNA from the original sample will not help. In such cases, it would be informative to examine the host species identity by sequencing a shorter region of the hosts cytochrome b or COI genes that can be amplified by any of many bird-universal primers (Kocher et al., 1989).

The examples we report here of published records that are possibly the result of PCR contamination have been singled out from cases where we have robust information on the natural occurrence of the parasite (host species and/or geographic region). We expect that these are just the tip of the problem because for most of the >3,900 lineages found to date, we know far too little of haemosporidian natural occurrence to be alerted if a record is improbable. Errors assigning parasites to host species is a problem that goes beyond the study itself, because when entered to common databases, the records will repeat the errors in global analyses of parasite host and distribution ranges (Clark et al., 2014). A relevant question to ask is whether PCR contaminations in publications, and as a consequence falsely assigned host species, have contributed to the general conclusion that haemosporidian parasites mainly evolve by frequent host switching rather than co-speciation (Fecchio et al., 2018). We think this is unlikely because switching between host species (Ricklefs et al., 2014) and continents (Ellis et al., 2019) have also been inferred from records of closely related but yet different lineages, and because they have different sequences, they cannot be caused by PCR-contaminations.

We believe that contamination of samples during PCR analysis is not exclusive to haemosporidian research, but an example of a broader problem that applies to molecular identification of other taxonomic groups – notably microorganisms - where opportunities to confirm the findings with alternative methods are restricted. Particularly challenging are metabarcoding studies where identifications are based only on DNA (Bush et al., 2019). For example, contamination, PCR errors and hopping index primers in multiplexed sequence pools will contribute to an omnipresence of low levels of false positives (Ficetola et al., 2015; van der Valk et al., 2020). Repeated analyses of re-extracted samples combined with careful bioinformatic filtering are hence required steps to minimize the presence of false positives (Ficetola et al., 2015). However, the optimal level of replication is always study-specific and as for all studies based on PCR amplification of low concentration or degraded DNA, the recommendation to "Look before you leap" (Taberlet et al., 1999) still holds true.

In conclusion, we hope that careful laboratory routines, including following up the presence of amplifications of negative controls, selection of positive controls, and verification of suspicious findings with alternative primers will reduce the publication of contamination records in the future. Accordingly, we encourage reviewers and journal editors to require that unexpected findings should be tested with additional primers before recommending manuscripts for publication.

AUTHOR CONTRIBUTIONS

S.B. conceived the study and wrote the article. All authors read and commented on the draft manuscript and approved final version.

ORCID

Staffan Bensch 🔟 https://orcid.org/0000-0002-0082-0899

REFERENCES

Aurrecoechea, C., Barreto, A., Basenko, E. Y., Brestelli, J., Brunk, B. P., Cade, S., Crouch, K., Doherty, R., Falke, D., Fischer, S., Gajria, B., Harb, O. S., Heiges, M., Hertz-Fowler, C., Hu, S., Iodice, J., Kissinger, J. C., Lawrence, C., Li, W., … Zheng, J. (2017). EuPathDB: The eukaryotic pathogen genomics database resource. *Nucleic Acids Research*, *45*, D581–D591. https://doi.org/10.1093/nar/gkw1105

Bensch, S., & Hellgren, O. (2020). The use of molecular methods in studies of avian Haemosporidians. In D. Santiago-Alarcón, & A. Marzal (Eds.), *Avian Malaria and Related Parasites in the Tropics: Ecology, Evolution, and Systematics*. (pp. 113-135). Springer. ISBN 978-3-030-51632-1.

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

Bensch, S., Stjernman, M., Hasselquist, D., Östman, Ö., Hansson, B., Westerdahl, H., & Torres-Pinheiro, R. (2000). Host specificity in avian blood parasites: a study of Plasmodium and Haemoproteus mitochondrial DNA amplified from birds. *Proceedings of the Royal Society of London B, Biological Sciences*, *267*, 1583–1589.

Bonneaud, C., Pérez-Tris, J., Federici, P., Chastel, O., & Sorci, G. (2006). Major histocompatibility alleles associated with local resistance to malaria in a passerine. *Evolution*, *60*, 383–389. https://doi.org/10.1111/j.0014-3820.2006.tb01114.x

Bush, A., Compson, Z. G., Monk, W. A., Porter, T. M., Steeves, R., Emilson, E., Gagne, N., Hajibabaei, M., Roy, M., & Baird, D. J. (2019). Studying ecosystems with DNA metabarcoding: Lessons from biomonitoring of aquatic macroinvertebrates. *Frontiers in Ecology and Evolution*, *7*, 434. https://doi.org/10.3389/fevo.2019.00434

Casanueva, P., Fernández, M., Rojo, M. À., & Campos, F. (2012). High prevalence of haemosporidian parasites infection in southern grey shrike *Lanius meridionalis* (Laniidae, Aves) from agricultural areas. *Italian Journal of Zoology*, *79*, 315–318.

Ciloglu, A., Ellis, V. A., Bernotiene, R., Valkiūnas, G., & Bensch, S. (2019). A new one-step multiplex PCR assay for simultaneous detection and identification of avian haemosporidian parasites. *Parasitology Research*, *118*, 191–201. https://doi.org/10.1007/s00436-018-6153-7

Ciloglu, A., Ellis, V. A., Duc, M., Downing, P. A., Inci, A., & Bensch, S. (2020). Evolution of vector transmitted parasites by host switching revealed through sequencing of *Haemoproteus* parasite mitochondrial genomes. *Molecular Phylogenetics and Evolution*, *153*, 106947. https://doi.org/10.1016/j.ympev.2020.106947

Clark, N. J., Clegg, S. M., & Lima, M. R. (2014). A review of global diversity in avian haemosporidians (*Plasmodium* and *Haemoproteus*: Haemosporida): New insights from molecular data. *International Journal for Parasitology*, *44*, 329–338. https://doi.org/10.1016/j.ijpara.2014.01.004

Dimitrov, D., Zehtindjiev, P., & Bensch, S. (2010). Genetic diversity of avian blood parasites in SE Europe: Cytochrome b lineages of the genera *Plasmodium* and *Haemoproteus* (Haemosporida) from Bulgaria. *Acta Parasitologica*, *55*, 201–209. https://doi.org/10.2478/s11686-010-0029-z

Dunn, J. C., Goodman, S. J., Benton, T. G., & Hamer, K. C. (2014). Active blood parasite infection is not limited to the breeding season in a declining farmland bird. *Journal of Parasitology*, *100*, 260–266. https://doi.org/10.1645/13-256.1

Ellis, V. A., Sari, E. H. R., Rubenstein, D. R., Dickerson, R. C., Bensch, S., & Ricklefs, R. E. (2019). The global biogeography of avian haemosporidian parasites is characterized by local diversification and intercontinental dispersal. *Parasitology*, *146*, 213–219. https://doi.org/10.1017/S0031182018001130

Escalante, A. A., Freeland, D. E., Collins, W. E., & Lal, A. A. (1998). The evolution of primate malaria parasites based on the gene encoding cytochrome b from the linear mitochondrial genome. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 8124–8129. https://doi.org/10.1073/pnas.95.14.8124

Fallon, S. M., Ricklefs, R. E., Swanson, B. L., & Bermingham, E. (2003). Detecting avian malaria: An improved polymerase chain reaction diagnostic. *Journal of Parasitology*, *89*, 1044–1047. https://doi.org/10.1645/GE-3157

Fecchio, A., Bell, J. A., Collins, M. D., Farias, I. P., Trisos, C. H., Tobias, J. A., Tkach, V. V., Weckstein, J. D., Ricklefs, R. E., & Batalha-Filho, H. (2018). Diversification by host switching and dispersal shaped the diversity and distribution of avian malaria parasites in Amazonia. *Oikos*, *127*, 1233–1242. https://doi.org/10.1111/oik.05115

Fernández, M., Rojo, M. À., Casanueva, P., Carrión, S., Hernández, M. Á., & Campos, F. (2010). High prevalence of haemosporidians in reed warbler *Acrocephalus scirpaceus* and sedge warbler *Acrocephalus schoenobaenus* in Spain. *Journal of Ornithology*, *151*, 27–32. https://doi.org/10.1007/s10336-009-0417-z

Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., Gielly, L., Lopes, C. M., Boyer, F., Pompanon, F., Rayé, G., & Taberlet, P. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*, *15*, 543–556. https://doi.org/10.1111/1755-0998.12338

Hellgren, O., Krizanauskiene, A., Hasselquist, D., & Bensch, S. (2011). Low haemosporidian diversity and one key-host species in a bird malaria community on a mid-Atlantic island (Sao Miguel, Azores). *Journal of Wildlife Diseases*, *47*, 849–859.

Hellgren, O., Waldenström, J., & Bensch, S. (2004). A new PCR assay for simultaneous studies of *Leucocytozoon*, *Plasmodium*, and *Haemoproteus* from avian blood. *Journal of Parasitology*, *90*, 797–802. https://doi.org/10.1645/GE-184R1

Hernández, M. Á., Rojo, M. Á., Campos, F., Gutiérrez-Corchero, F., & Moreno-Rueda, G. (2017). Haemosporidian prevalence in Southern Grey Shrike *Lanius meridionalis* nestlings: Impact on body condition and geographic distribution in the Iberian Peninsula. *Bird Study*, *64*, 362–373. https://doi.org/10.1080/00063657.2017.1364219

Huang, X., Hansson, R., Palinauskas, V., Valkiūnas, G., Hellgren, O., & Bensch, S. (2018). The success of sequence capture in relation to phylogenetic distance from a reference genome: A case study of avian haemosporidian parasites. *International Journal for Parasitology*, *48*, 947–954. https://doi.org/10.1016/j.ijpara.2018.05.009

Imura, T., Suzuki, Y., Ejiri, H., Sato, Y., Ishida, K., Sumiyama, D., Murata, K., & Yukawa, M. (2011). Prevalence of avian haematozoa in wild birds in a high-altitude forest in Japan. *Veterinary Parasitology*, *183*, 244–248. https://doi.org/10.1016/j.vetpar.2011.07.027

Ishtiaq, F., Gering, E., Rappole, J. H., Rahmani, A. R., Jhala, Y. V., Dove, C. J., Milensky, C., Olson, S. L., Peirce, M. A., & Fleischer, R. C. (2007). Prevalence and diversity of avian hematozoan parasites in Asia: A regional survey. *Journal of Wildlife Diseases*, *43*, 382–398. https://doi.org/10.7589/0090-3558-43.3.382

Jarvis, E. D., Mirarab, S., Aberer, A. J., Li, B., Houde, P., Li, C., Ho, S. Y. W., Faircloth, B. C., Nabholz, B., Howard, J. T., Suh, A., Weber, C. C., da Fonseca, R. R., Li, J., Zhang, F., Li, H., Zhou, L., Narula, N., Liu, L., ... Zhang, G. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*, *346*, 1320–1331. https://doi.org/10.1126/science.1253451

Jolley, K. A., Bray, J. E., & Maiden, M. C. J. (2018). Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications [version 1; peer review: 2 approved]. *Wellcome Open Research*, *3*, 124.

Kim, K. S., & Tsuda, Y. (2010). Seasonal changes in the feeding pattern of *Culex pipiens pallens* govern the transmission dynamics of multiple lineages of avian malaria parasites in Japanese wild bird community. *Molecular Ecology*, *19*, 5545–5554. https://doi.org/10.1111/j.1365-294X.2010.04897.x

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X., & Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the United States of America*, *86*, 6196–6200. https://doi.org/10.1073/pnas.86.16.6196

Krizanauskiene, A., Iezhova, T. A., Palinauskas, V., Chernetsov, N., & Valkiūnas, G. (2012). *Haemoproteus nucleocondensus* n. sp. (Haemosporida, Haemoproteidae) from a Eurasian songbird, the Great Reed Warbler *Acrocephalus arundinaceus*. *Zootaxa*, *3441*, 36–46.

Lacorte, G. A., Félix, G. M. F., Pinheiro, R. R. B., Chaves, A. V., Almeida-Neto, G., Neves, F. S., Leite, L. O., Santos, F. R., & Braga, É. M. (2013). Exploring the diversity and distribution of Neotropical avian malaria parasites - a molecular survey from southeast Brazil. *PLoS One*, *8*, e57770. https://doi.org/10.1371/journal.pone.0057770

Levin, I. I., Valkiūnas, G., Santiago-Alarcon, D., Cruz, L. L., Iezhova, T. A., O'Brien, S. L., Hailer, F., Dearborn, D., Schreiber, E. A., Fleischer, R. C., Ricklefs, R. E., & Parker, P. G. (2011). Hippoboscid-transmitted *Haemoproteus* parasites (Haemosporida) infect Galapagos Pelecaniform birds: Evidence from molecular and morphological studies, with a description of *Haemoproteus iwa*. *International Journal for Parasitology*, *41*, 1019–1027. https://doi.org/10.1016/j.ijpara.2011.03.014

Marzal, A., Garcia-Longoria, L., Callirgos, J. M. C., & Sehgal, R. N. M. (2015). Invasive avian malaria as an emerging parasitic disease in native birds of Peru. *Biological Invasions*, *17*, 39–45. https://doi.org/10.1007/s10530-014-0718-x

Mata, V. A., da Silva, L. P., Lopes, R. J., & Drovetski, S. V. (2015). The Strait of Gibraltar poses an effective barrier to host-specialised but not to host-generalised lineages of avian Haemosporidia. *International Journal for Parasitology*, *45*, 711–719. https://doi.org/10.1016/j.ijpara.2015.04.006

Moens, M. A. J., Valkiūnas, G., Paca, A., Bonaccorso, E., Aguirre, N., & Pérez-Tris, J. (2016). Parasite specialization in a unique habitat: Hummingbirds as reservoirs of generalist blood parasites of Andean birds. *Journal of Animal Ecology*, *85*, 1234–1245. https://doi.org/10.1111/1365-2656.12550

Musa, S., Fachet, K., Dinkel, A., Mackenstedt, U., & Woog, F. (2018). Low prevalence of haemosporidian parasites in resident Greylag Geese (*Anser anser*) in southwestern Germany. *Journal of Ornithology*, *159*, 1099–1103. https://doi.org/10.1007/s10336-018-1580-x

Musa, S., Mackenstedt, U., Woog, F., & Dinkel, A. (2019). Avian malaria on Madagascar: Prevalence, biodiversity and specialization of haemosporidian parasites. *International Journal for Parasitology*, *49*, 199–210. https://doi.org/10.1016/j.ijpara.2018.11.001

Nourani, L., Aliabadian, M., Mirshamsi, O., & Djadid, N. D. (2018). Molecular detection and genetic diversity of avian haemosporidian parasites in Iran. *PLoS One*, *13*, e0206638. https://doi.org/10.1371/journal.pone.0206638

Pacheco, M. A., Matta, N. E., Valkiūnas, G., Parker, P. G., Mello, B., Stanley, C. E., Lentino, M., Garcia-Amado, M. A., Cranfield, M., Kosakovsky Pond, S. L., & Escalante, A. A. (2018). Mode and rate of evolution of Haemosporidian mitochondrial genomes: Timing the radiation of avian parasites. *Molecular Biology and Evolution*, *35*, 383–403. https://doi.org/10.1093/molbev/msx285

Palinauskas, V., Iezhova, T. A., Krizanauskiene, A., Markovets, M. Y., Bensch, S., & Valkiūnas, G. (2013). Molecular characterization and distribution of *Haemoproteus minutus* (Haemosporida, Haemoproteidae): A pathogenic avian parasite. *Parasitology International*, *62*, 358–363. https://doi.org/10.1016/j.parint.2013.03.006

Perkins, S. L., & Schall, J. J. (2002). A molecular phylogeny of malarial parasites recovered from cytochrome b gene sequences. *Journal of Parasitology*, *88*, 972–978.

Ricklefs, R. E., Outlaw, D. C., Svensson-Coelho, M., Medeiros, M. C. I., Ellis, V. A., & Latta, S. (2014). Species formation by host shifting in avian malaria parasites. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 14816–14821. https://doi.org/10.1073/pnas.1416356111

Rivero, A., & Gandon, S. (2018). Evolutionary ecology of avian malaria: Past to present. *Trends in Parasitology*, *34*, 712–726. https://doi.org/10.1016/j.pt.2018.06.002

Rojo, M. À., Hernández, M. À., Campos, F., Santamaría, T., Dias, S., & Casanueva, P. (2015). The Iberian Peninsula is an area of infection by *Haemoproteus payevskyi* and *Haemoproteus nucleocondensus* for the white-throated dipper *Cinclus cinclus*. *Ardeola*, *62*, 373–382.

Sehgal, R. N. M. (2019). Recognition of the 4th International Conference on Malaria and Related Haemosporidian Parasites of Wildlife, held Nov. 1–5, 2018 in Beijing China. *International Journal for Parasitology-Parasites and Wildlife*, *8*, 33–35.

Stjernman, M., Råberg, L., & Nilsson, J. A. (2008). Maximum host survival at intermediate parasite infection intensities. *PLoS One*, *3*, e2463. https://doi.org/10.1371/journal.pone.0002463

Taberlet, P., Waits, L. P., & Luikart, G. (1999). Noninvasive genetic sampling: Look before you leap. *Trends in Ecology and Evolution*, *14*, 323–327. https://doi.org/10.1016/S0169-5347(99)01637-7

Turcotte, A., Belisle, M., Pelletier, F., & Garant, D. (2018). Environmental determinants of haemosporidian parasite prevalence in a declining population of Tree swallows. *Parasitology*, *145*, 961–970. https://doi.org/10.1017/S0031182017002128

Valkiūnas, G. (2005). *Avian malaria parasites and other Haemosporidia*, Boca Raton: CRC Press. ISBN 0-415-30097-5.

Valkiūnas, G., Krizanauskiene, A., Iezhova, T. A., Hellgren, O., & Bensch, S. (2007). Molecular phylogenetic analysis of circumnuclear hemoproteids (Haemosporida : Haemoproteidae) of sylviid birds, with a description of *Haemoproteus parabelopolskyi* sp nov. *Journal of Parasitology*, *93*, 680–687. https://doi.org/10.1645/GE-1102R.1

Valkiūnas, G., Zehtindjiev, P., Hellgren, O., Ilieva, M., Iezhova, T. A., & Bensch, S. (2007). Linkage between mitochondrial cytochrome b lineages and morphospecies of two avian malaria parasites, with a description of *Plasmodium* (*Novyella*) *ashfordi* sp. nov. *Parasitology*

Research, *100*, 1311–1322. https://doi.org/10.1007/s0043 6-006-0409-3

van der Valk, T., Vezzi, F., Ormestad, M., Dalén, L., & Guschanski, K. (2020). Index hopping on the Illumina HiseqX platform and its consequences for ancient DNA studies. *Molecular Ecology Resources*, *20*, 1171–1181. https://doi.org/10.1111/1755-0998.13009