

NEWS AND VIEWS

Perspective

Promising protocols for parasites: Metatranscriptomics improves detection of hyperdiverse but low abundance communities

Loren Cassin-Sackett 

Department of Biology, University of Louisiana, Lafayette, LA, USA

Correspondence

Loren Cassin-Sackett, Department of Biology, University of Louisiana, Lafayette, LA, USA.

Email: cassin.sackett@gmail.com

Genomic technologies continue to shed light on important ecological and evolutionary questions. Nonetheless, these new tools are applied disproportionately in a small fraction of global biodiversity, partly because of technical challenges to studying highly diverse taxa that occur in low abundances in an environment (e.g., marine and microbial communities). As a result, our understanding of ecological and evolutionary processes lags in many taxa. In a From the Cover manuscript in this issue of *Molecular Ecology Resources*, Galen, Borner, Williamson, Witt, and Perkins (2020) present a novel approach for characterizing diversity that combines metatranscriptomics with rigorous bioinformatic processing to dramatically improve detection and identification of diverse, low-abundance avian blood parasites. Their approach is an exciting application of available tools that increases our potential for a deeper understanding of diversity in other communities of low-abundance, highly diverse taxa.

KEYWORDSbioinformatics, environmental DNA, infection intensity, *Plasmodium*, species identification

Parasites comprise an astonishing proportion of global biodiversity, yet a minute fraction of genetically characterized taxa, leading to a vast underrepresentation of parasite diversity in our sequence repositories. The paucity of parasite sequence data can be traced to several technical difficulties in obtaining such data, including the swamping by host genetic material, multiple infections within a host, and extreme morphological similarity among species that lead to undescribed cryptic diversity.

Traditionally, scientists studied blood parasite species diversity using morphological traits, but mitochondrial sequencing revealed widespread morphological convergence and phenotypic plasticity (Perkins, Martinsen, & Falk, 2011). Since then, cytochrome b genotypes have been the standard for detecting and identifying most blood parasites. However, simultaneously amplifying highly divergent taxa (especially when diversity is unknown a priori) is challenging and labour intensive (Pacheco et al., 2018). Moreover, technical challenges have hindered the ability of blood parasite research to

capitalize on many advances in genomic technologies. For instance, incidental capture of host DNA is challenging to avoid, thus necessitating higher coverage sequencing to compensate for numerous off-target reads. Additionally, most library preparation protocols are not optimized for the AT-rich genomes that characterize blood parasites, leading to poor recovery (Oyola et al., 2012). As a result, our inference of the genetic and species diversity of blood parasites is underestimated, with the underrepresentation in genomic repositories especially striking for avian blood parasites (Figure 1).

Galen et al. (2020) crafted an innovative way of detecting low-intensity infections and assessing diversity with higher precision than microscopy-based methods or mitochondrial genotyping. The authors assembled transcriptomes and separated parasite from host transcripts by selecting contigs with the best BLAST match to apicomplexan proteomes. Sequences were then assigned to parasite genus (*Plasmodium*, *Parahaemoproteus*, or *Leucocytozoon*) using a novel resampling strategy: For each target contig, a single sequence was randomly sampled from each of the three reference genera (thus creating a “quartet” of sequences) and pairwise sequence identity

LCS performed the analysis, wrote, and edited the manuscript.

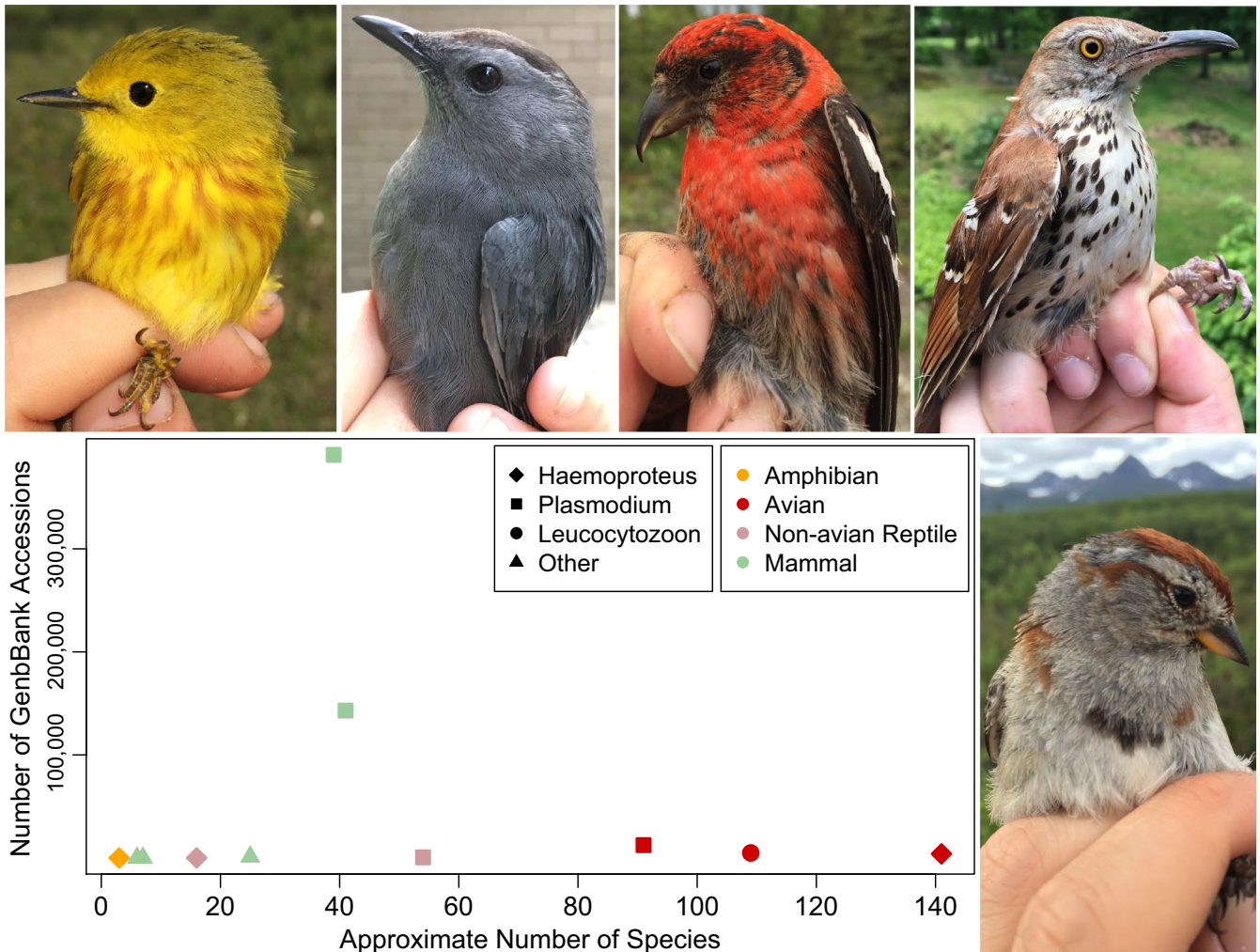


FIGURE 1 Clockwise, from top left: Examples of avian species sampled in this study - Yellow warbler (*Setophaga petechia*), Gray catbird (*Dumetella carolinensis*), White-winged crossbill (*Loxia leucoptera*), Brown thrasher (*Toxostoma rufum*), and American tree sparrow (*Spizelloides arborea*). Bottom left - The number of GenBank accessions available for each group of blood parasites, relative to the approximate number of species in that group. By comparison, the bacterial genus *Escherichia* contains ~6 species and over 7 million GenBank accessions. Catbird image from Loren Cassin Sackett; all others courtesy Spencer Galen

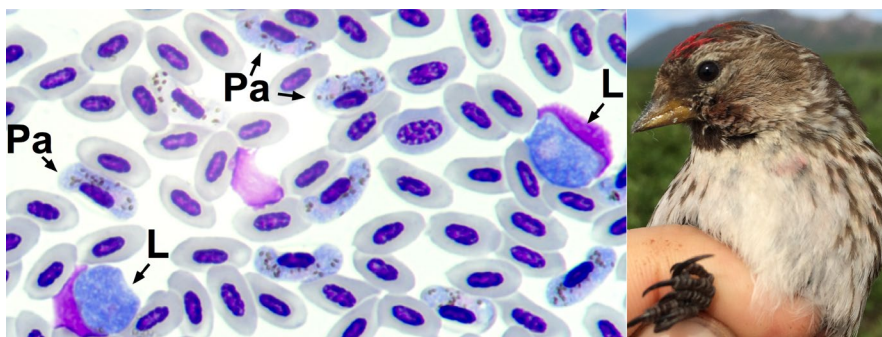


FIGURE 2 A Common redpoll (*Acanthis flammea*, right) coinfecting with both *Parahaemoproteus* ("Pa" in microscopy image, left) and *Leucocytozoon* ("L")

was estimated to choose the closest match. This procedure was repeated for all possible combinations of one target contig + three reference genus sequences. This approach was successful at distinguishing parasite identity even in cases of coinfection (Figure 2) of up to six parasite lineages. Although largely effective, this approach may be subject to certain types of errors such as incorrect assignment or the false inference of isoform variants as distinct sequences,

thus warranting subsequent validation steps. Another drawback is that infection intensity influenced both the number of transcripts and the set of genes recovered. For very low intensity infections, the genes recovered were a relatively small subset of the total, which could make some downstream analyses more difficult (although traditional methods also suffer problems in low-intensity infections). Metatranscriptomics is also costly (e.g., sequencing costs alone were

likely ~\$250/sample in this study), although this cost is constantly declining.

Nonetheless, this approach markedly improves our ability to study Haemosporidian parasite diversity. The method detected nearly 25% more infections than either microscopy or mtDNA amplification (Galen et al., 2020), proving particularly effective in very low-intensity infections (e.g., 0.03% parasitaemia). The method also offers advantages over other genomic tools: for instance, although DNA sequence capture can also detect low-intensity infections, performance declines as relatedness among taxa decreases (Huang et al., 2018), leading to similar biases as those occurring with traditional PCR amplification (Valkiūnas et al., 2006), such as underestimates of coinfection and genetic variability. In contrast, metatranscriptomics is unconstrained by taxonomic divergence.

The approach also offers promise for using functional diversity to delineate blood parasite species in a more cohesive way than the presently used combination of morphology and cytochrome b variation. At appropriate sequencing depth, metatranscriptomics can simultaneously provide data on both gene expression and sequence variation, allowing us to probe questions in disease ecology and evolution that have been elusive, such as: Which genes are the most divergent across blood parasites, and are these implicated in species divergence (Pérez-Tris et al., 2007)? What are the genes expressed by parasite taxa infecting different hosts (Videvall et al., 2017), and which genes govern interactions among parasite species versus interactions with the host? How many parasite species can coexist in a particular host tissue, and how does phylogenetic distance influence coinfection (Galen, Speer, & Perkins, 2019)? How stable are coinfections, what host and parasite conditions favour coinfection, and which taxa are most likely to coinfect? In conjunction with other methods such as experimental infections, this approach can shed light on important ecological and evolutionary questions.

The approach presented by Galen et al. (2020) is not without drawbacks, the largest being cost and extensive computational labour. Nonetheless, the combination of metatranscriptomics and bioinformatic assignment to parasite genus provides markedly better estimates of parasite diversity than the currently used methods due to its ability to detect and identify infections at very low intensity. The approach will be especially useful in systems with low DNA concentration or in which the DNA of the host (or one dominant species) swamps that of the parasites (or other species in the community). For instance, the combined metatranscriptomic and bioinformatic assignment approach could be applied to other types of multiparasite communities, to low-abundance, high diversity communities such as those found in marine environments, and other types of environmental DNA containing poorly described taxa. This method

advances our ability to characterize diversity across underrepresented branches of the tree of life.

ORCID

Loren Cassin-Sackett  <https://orcid.org/0000-0002-6000-4789>

REFERENCES

- Galen, S., Borner, J., Williamson, J., Witt, C., & Perkins, S. (2020). Metatranscriptomics yields new genomic resources and sensitive detection of infections for diverse blood parasites. *Molecular Ecology Resources*, 20.
- Galen, S. C., Speer, K. A., & Perkins, S. L. (2019). Evolutionary lability of host associations promotes phylogenetic overdispersion of co-infecting blood parasites. *Journal of Animal Ecology*, 88, 1936–1949. <https://doi.org/10.1111/1365-2656.13089>
- Huang, X. I., 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 of Parasitology*, 48(12), 947–954. <https://doi.org/10.1016/j.ijpara.2018.05.009>
- Oyola, S. O., Otto, T. D., Gu, Y., Maslen, G., Manske, M., Campino, S., ... Quail, M. A. (2012). Optimizing Illumina next-generation sequencing library preparation for extremely AT-biased genomes. *BMC Genomics*, 13, 1. <https://doi.org/10.1186/1471-2164-13-1>
- Pacheco, M. A., Cepeda, A. S., Bernotienė, R., Lotta, I. A., Matta, N. E., Valkiūnas, G., & Escalante, A. A. (2018). Primers targeting mitochondrial genes of avian haemosporidians: PCR detection and differential DNA amplification of parasites belonging to different genera. *International Journal of Parasitology*, 48(8), 657–670. <https://doi.org/10.1016/j.ijpara.2018.02.003>
- Pérez-Tris, J., Hellgren, O., Križanauskienė, A., Waldenström, J., Secondi, J., Bonneaud, C., ... Bensch, S. (2007). Within-host speciation of malaria parasites. *PLoS ONE*, 2(2), e235. <https://doi.org/10.1371/journal.pone.0000235>
- Perkins, S. L., Martinsen, E. S., & Falk, B. G. (2011). Do molecules matter more than morphology? *Promises and Pitfalls in Parasites. Parasitology*, 138(13), 1664–1674. <https://doi.org/10.1017/S0031182011000679>
- Valkiūnas, G., Bensch, S., Iezhova, T. A., Križanauskienė, A., Hellgren, O., & Bolshakov, C. V. (2006). Nested cytochrome b polymerase chain reaction diagnostics underestimate mixed infections of avian blood haemosporidian parasites: Microscopy is still essential. *Journal of Parasitology*, 92(2), 418–422.
- Videvall, E., Cornwallis, C. K., Ahrén, D., Palinauskas, V., Valkiūnas, G., & Hellgren, O. (2017). The transcriptome of the avian malaria parasite *Plasmodium ashfordi* displays host-specific gene expression. *Molecular Ecology*, 26(11), 2939–2958.

How to cite this article: Cassin-Sackett L. Promising protocols for parasites: Metatranscriptomics improves detection of hyperdiverse but low abundance communities. *Mol Ecol Resour.* 2019;00:1–3. <https://doi.org/10.1111/1755-0998.13122>