International Journal for Parasitology 55 (2025) 657–663

Contents lists available at ScienceDirect

International Journal for Parasitology

journal homepage: www.elsevier.com/locate/ijpara

Biased availability of genetic data for parasites: it's all about the host ☆

Chen-Hua Li *, Robert Poulin

Department of Zoology, University of Otago, PO Box 56, Dunedin, New Zealand

ARTICLE INFO

Article history:
Received 20 January 2025
Received in revised form 27 May 2025
Accepted 28 May 2025
Available online 31 May 2025

Keywords:
Genetic information
DNA sequencing
Nucleotide sequences
Phylogenetic reconstruction
Parasite evolution
Helminth

ABSTRACT

Advances in DNA sequencing technology have significantly increased the availability of publicly accessible genetic data across various parasite taxa. This genetic data is crucial for elucidating gene flow, connectivity among parasite populations, cryptic diversity, and resolving parasite phylogenies. However, phylogenetic reconstructions are often hindered by incomplete genetic data across taxa, particularly in less-studied taxa like parasitic helminths. We tested whether the availability of genetic data for helminth species is biased and influenced by which hosts they infect using the largest known helminth life cycle database. We compared helminth species with genetic sequences in the NCBI Nucleotide and Gene databases against those without publicly available sequences. We examined the impact of the number of definitive hosts, their higher taxon, conservation status, and habitat on genetic data availability. Our findings reveal significant biases in existing genetic data for helminth parasites, influenced by host-related factors. Helminth species with more definitive hosts species, hosts of conservation concerns, and/or those with terrestrial hosts are more likely to have genetic data available. These biases in genetic data availability raise concerns for phylogenetic studies, as they suggest that the current genetic knowledge of helminth parasites is neither random nor representative of existing biodiversity. Consequently, phylogenetic trees based on biased data may not accurately capture the true evolutionary relationships among parasite taxa, as well as trends in the evolution of key traits, such as host specificity. Comprehensive and unbiased data collection efforts are needed to improve the accuracy of phylogenetic analyses and our understanding of parasite evolution.

1. Introduction

Over the past two decades, the rapid drop in the cost of DNA sequencing (see National Human Genome Research Institute, USA; https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data) and the growing incorporation of genetic data in parasite taxonomy and systematics (Perkins et al., 2011; Blasco-Costa et al., 2016) have resulted in an increasing number of publicly available DNA sequences spanning all parasite taxa (Selbach et al., 2019). Beyond taxonomy and systematics, the availability of DNA sequences has proven essential to elucidate the extent of gene flow and connectivity among parasite populations (Blasco-Costa and Poulin, 2013; Thorn et al., 2023) and to uncover cryptic parasite diversity (Nadler and Pérez-Ponce de León, 2011; Pérez-Ponce de León and Poulin, 2018). On a larger evolutionary timescale, genetic data is also crucial to resolve parasite phylogenies and host–parasite coevolutionary history (e.g., Olson and Tkach, 2005; Caira et al., 2014; Pérez-Ponce de León and Hernández-Mena, 2019).

Phylogenetic reconstructions of the evolutionary history of any higher taxon are only as reliable as the taxonomic completeness of the genetic data on which they are based, however. Firstly, for most extant species there isn't any genetic data available in any public repository, let alone genomic data, which would allow further resolution. Even for well–studied taxa such as tetrapod vertebrates, nearly one–quarter of species lack genetic data (Šmíd, 2022). As a consequence, the placement of more than a third of tetrapod species within global phylogenetic trees has to be imputed, i.e. approximated from that of their closest relatives for which there is genetic data (Jetz and Pyron, 2018; Upham et al., 2019). This problem is likely much worse for less studied taxa, such as parasitic helminths. The unknown exact placement of most species within phylogenies has been termed the Darwinian shortfall (Diniz-Filho et al., 2013) and is seen as a major impediment for any evolutionary analysis of biodiversity.

Secondly, in addition to the Darwinian shortfall, what if available genetic data do not provide a random and representative sample of living species, but are instead consistently biased in some way? For instance, among the thousands of helminth species

☆ Note: Supplementary data associated with this article.

* Corresponding author.
E-mail address: cillialichen@hotmail.com (C.-H. Li).

https://doi.org/10.1016/j.ijpara.2025.05.007
0020-7519/

discovered and described in the past two decades, the probability that they are the subject of further research is certainly not equal, with various factors creating biases in research effort among species (Poulin et al., 2023). If nucleotide sequences are more likely to be obtained for species from certain geographic regions or with certain traits than for species from other regions or with different traits, then phylogenetic reconstructions using this biased data are unlikely to provide a true picture of parasite evolution, diversification or host–switching. For example, among the well–studied vertebrates, species for which genetic data are available are significantly larger-bodied and wider-ranging than those lacking genetic data (Guedes et al., 2024). Genetic information on parasitic helminths is likely to show biases of this sort, with serious consequences for large–scale phylogenetic analyses.

An earlier study has indeed identified geographic biases in the genetic study of parasites: the proportion of known helminth species for which genetic data are available is disproportionately high in some countries or regions, and disproportionately low in others (Poulin et al., 2019). Here, we test a different general hypothesis: whether or not DNA sequences are available for particular parasite species depends mainly on what host(s) they infect. Specifically, we test four predictions. First, we expect that the taxon of the parasite's definitive host will affect the probability that DNA sequences are obtained for that parasite. For instance, parasites of mammals are likely the focus of greater research than those of reptiles, given greater economic and cultural interest in the former group. Second, generalist parasites that can infect a wide range of definitive hosts are expected to have been sequenced more frequently than specialist parasites that infect very few hosts, simply because the chances that researchers study parasites must depend on how widespread and commonly encountered they are. Third, since vertebrate species classified as vulnerable or of even greater conservation concern generally attract more research (Trimble and van Aarde, 2010; Robertson and McKenzie, 2015), we might expect the same for their parasites (see Poulin et al., 2023). Therefore, we predict that the conservation status of a parasite's host(s) will determine how likely it is that nucleotide sequences have been obtained for that parasite. Fourth, the habitat where hosts live can influence their ease of study and relevance to society, and thus we might expect that parasites of terrestrial hosts may be more likely to have been studied and sequenced than those of freshwater or marine hosts.

We test these predictions by comparing helminth species for which sequence data are available in the NCBI's Nucleotide or Gene databases with helminth species for which no sequence is publicly available in either of these databases. Our analysis considers three large groups of helminths, acanthocephalans, cestodes and nematodes, leveraging information from the largest available database on helminth life cycles (Benesh et al., 2017). In addition to comparisons among helminth groups, we compare helminth species with respect to: (i) the higher taxon of their definitive hosts; (ii) how many definitive host species they have; (iii) whether or not at least one of these definitive hosts is of conservation concern; and (iv) the habitat where their definitive hosts live. Any effect of these four host-related characteristics on the likelihood that sequence data are available for helminth species would reveal one or more biases in our genetic knowledge of parasites, casting some doubts on how accurately our existing data capture the full extent of parasite biodiversity.

2. Methods

2.1. Helminth database

The database on helminth lifecycles we utilized included information on host species and habitat for each parasite species, as well as their higher taxonomic affiliation (Benesh et al., 2017). It was compiled from the primary literature and, importantly, it includes all host species reported for each parasite species (current as of 2017). For each host species, data on its higher vertebrate group (mammal, bird, herptile, or fish) was also available. Of the 973 species in the database, we included 959 species in our analysis, and thus we excluded the nine species that did not have definitive host species information, and five species with no vertebrate definitive host. For each parasite species, we recorded the number of known definitive hosts. Also, the vertebrate host group was assigned based on the majority of a parasite's definitive hosts; in the vast majority of cases, all host species belonged to the same high group (e.g., all birds, all fish, etc.).

2.2. IUCN Red List Data

To obtain IUCN (International Union for Conservation of Nature) Red List information for the host species, we downloaded the dataset on November 27, 2024, containing the Red List categories: Critically Endangered, Endangered, Vulnerable, Lower Risk: Conservation Dependent, and Near Threatened (IUCN, version 2024-2). After matching host species with the species in those Red List categories, each parasite was categorized as either having at least one definitive host categorized as Near Threatened or higher, or not having any host in these categories.

2.3. Genetic information

Genetic information for cestode, nematode, and acanthocephalan species was obtained from two NCBI (National Center for Biotechnology Information) databases: the Nucleotide database and the Gene database. The Nucleotide database (https://www.ncbi.nlm.nih.gov/nuccore) contains sequences (coding or non-coding, DNA or RNA, partial or complete genes) for a given species from multiple sources including GenBank, Reference Sequence (RefSeq), Third Party Annotation (TPA), and Protein Data Bank (PDB). The Gene database (https://www.ncbi.nlm.nih.gov/gene) focuses on genes and provides detailed gene-specific information, including sequence, map, expression, function, structure, and related sequences.

The accession IDs of available nucleotide sequences of the above parasite groups (cestode, nematode, and acanthocephalan) in the NCBI Nucleotide database were downloaded as of December 2, 2024. Taxonomic information was retrieved from accession IDs using the package taxonomizr (version 0.10.6) (Sherrill-Mix, 2024) in R (R Core Team, 2024). The sequence record of the above three parasite groups in the NCBI Gene database were downloaded as of November 27, 2024.

From the downloaded data, we categorized each parasite as (i) either having or not having any sequences (hereafter referred to as Nucleotide Sequences) available in the NCBI Nucleotide database, and (ii) either having or not having any sequences (hereafter referred to as Gene Sequences) available in the NCBI Gene database. The genes for which sequencing information was available in either the Nucleotide or Gene databases varied among helminth species, with the main genes being the small subunit ribosomal RNA (18S rRNA), the large subunit ribosomal RNA (28S rRNA), mitochondrial cytochrome *c* oxidase subunit 1 (*cox*1), mitochondrial 12S ribosomal RNA, and mitochondrial 16S ribosomal RNA.

2.4. Statistical analysis and data visualization

Logistic regression analysis was performed with the availability of Nucleotide Sequences as the binary (available or not) response variable. Explanatory variables included helminth group (acanthocephalan, cestode or nematode), vertebrate host group (mammal,

bird, herptile, or fish), the number of definitive host species, whether or not the parasite had at least one host species categorized as Near Threatened or higher, and habitat (freshwater, marine, or terrestrial). In order to assess the significance of the explanatory variables, we conducted ANOVA on the logistic regression model. Model selection was based on the BIC (Bayesian Information Criterion) value. A similar analysis was conducted using the availability of Gene Sequences as the response variable. Data visualization was performed using the ggplot2 package (version 3.5.1, Wickham 2016) in R.

3. Results

Of the 959 species included in this study, 502 species have available Nucleotide Sequences, and 113 species have available Gene Sequences in the NCBI databases. Additionally, 316 species have at least one host species that is categorized as Near Threatened or above on the IUCN Red List. The majority of parasite species in our analyses were nematodes, followed by cestodes and acanthocephalans.

Whether or not the parasite has Nucleotide Sequences available in the NCBI Nucleotide database was affected by multiple variables (see Table S1). Parasites with a larger number of vertebrate definitive host species were more likely to have Nucleotide Sequences available ($\chi^2(1) = 20.77$, $p < 0.001$) (Fig. 1). In addition, if a parasite had at least one host categorized as near threatened or above in the IUCN red list, it was significantly more likely to have Nucleotide Sequences ($\chi^2(1) = 8.70$, $p < 0.01$) (Fig. 2). Moreover, the availability of Nucleotide Sequences among parasites differed based on their vertebrate host group ($\chi^2(3) = 50.19$, $p < 0.001$), with parasites from mammals being the most likely to have Nucleotide Sequences available (Fig. 3). Our analysis also revealed an interaction between parasite group and habitat ($\chi^2(4) = 12.84$, $p = 0.01$) (Fig. 4).

In terms of the availability of Gene Sequences in the NCBI Gene database (See Table S2), similar to the Nucleotide Sequences availability, it was influenced by the number of vertebrate definitive host species ($\chi^2(1) = 6.83$, $p < 0.01$) (Fig. 1) and vertebrate host group ($\chi^2(3) = 51{,}25$, $p < 0.001$) (Fig. 3), also with an interaction between parasite group and habitat ($\chi^2(4) = 12.63$, $p = 0.01$) (Fig. 4). Moreover, the habitat in which the parasite occurred influenced whether or not Gene Sequences was available ($\chi^2(2) = 6.67$, $p = 0.04$), with parasites from terrestrial habitats being the most likely to have Gene Sequences available (Fig. 5). However, the host IUCN status did not influence the Gene Sequences availability ($\chi^2(1) = 0.38$, $p = 0.53$) (Fig. 2).

4. Discussion

Knowledge of parasite biodiversity grows as more and more species are discovered, described and named (Poulin and Presswell, 2016). Genetic information on these species is lagging behind but catching up, as an increasing number of species are characterized genetically. This information is crucial for reconstructing the evolutionary history of parasites and their coevolutionary interactions with hosts, but also for their conservation (Velázquez-Urrieta et al., 2024). Even if genetic data are not available for every single known species, they can still yield valuable information if they are representative of the existing biodiversity. At this point in time, we can ask whether genetic information is

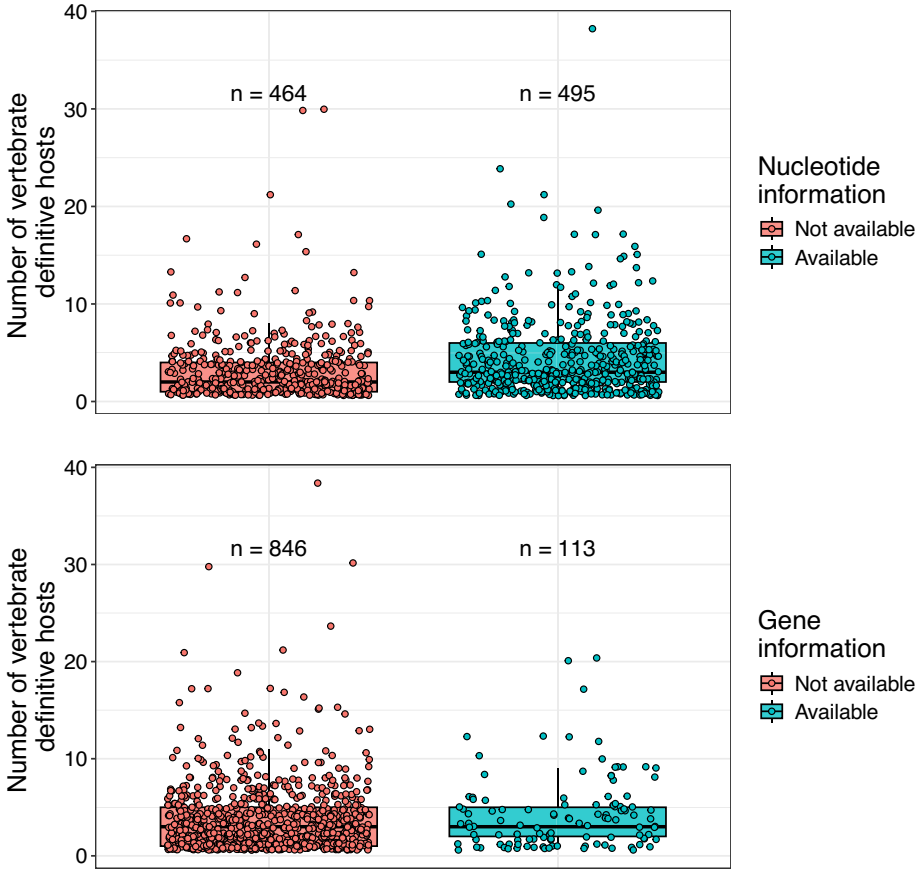

Fig. 1. Genetic data availability by number of hosts. Boxplot combined with scatter plot of number of vertebrate definitive hosts in helminth parasites for which genetic information is available in either the Nucleotide database (top) or Gene database (bottom) in NCBI. "n" represents sample size.

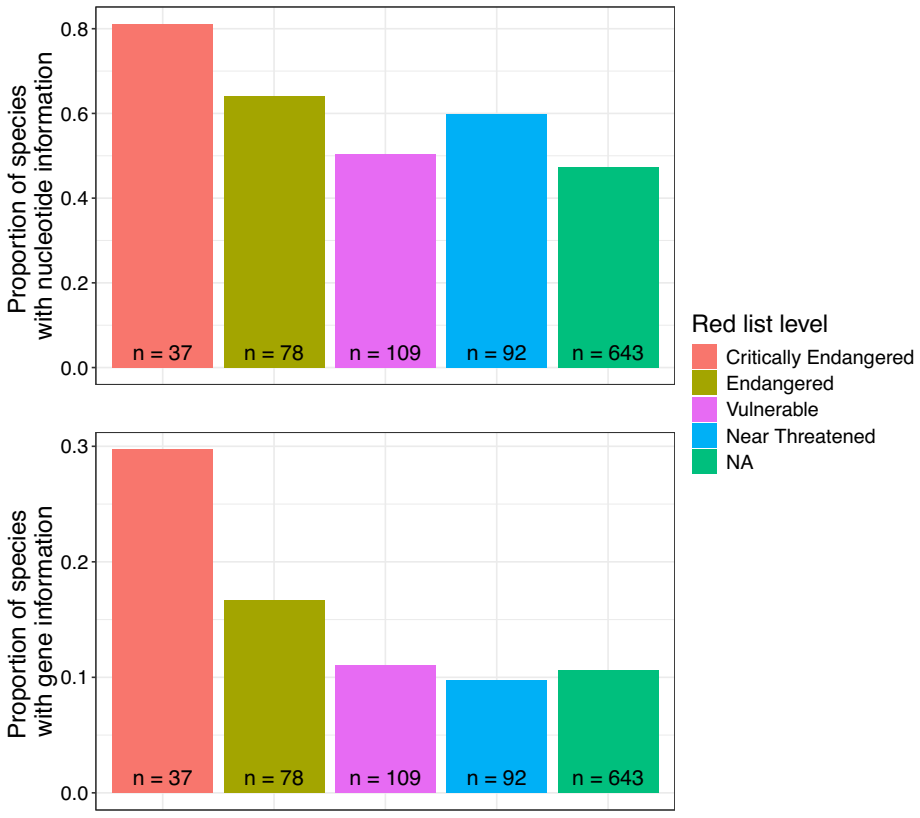

Fig. 2. Genetic data availability by host conservation status. Proportion of helminth parasite species for which genetic information is available either in the NCBI Nucleotide database (top) or Gene database (bottom) as a function of their host with the highest IUCN red list levels (critically endangered, endangered, vulnerable, near threatened, or NA). "n" represents the sample size. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

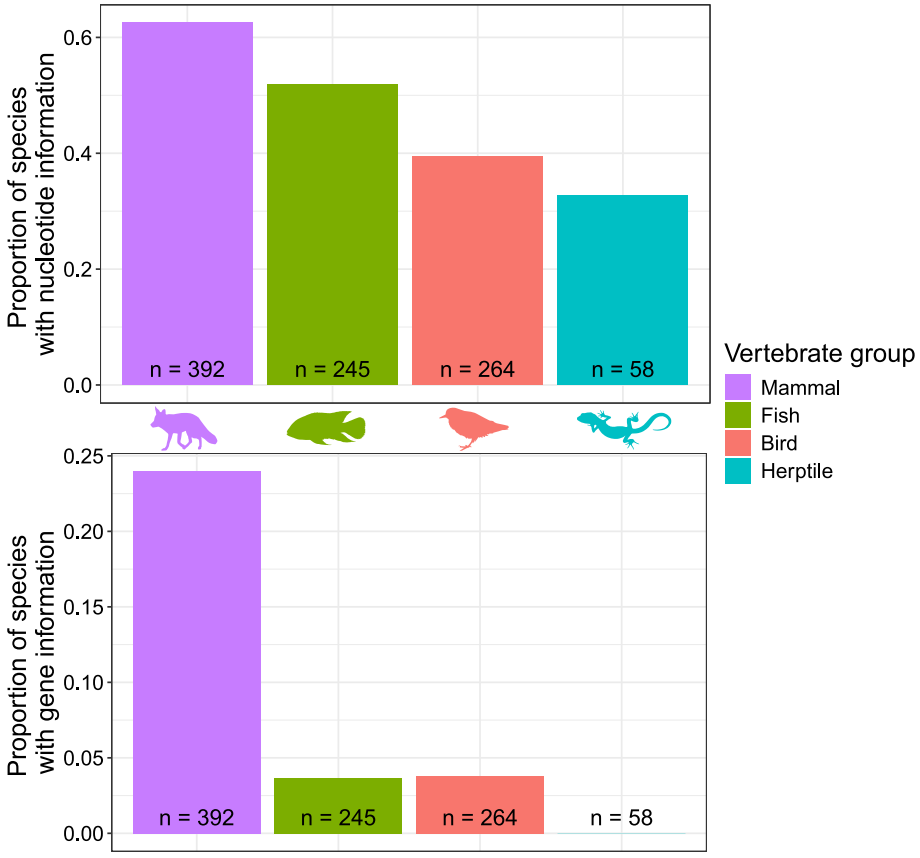

Fig. 3. Genetic data availability by host taxonomic group. Proportion of helminth parasite species for which genetic information is available either in the NCBI Nucleotide database (top) or Gene database (bottom) as a function of the taxonomic group of their vertebrate definitive hosts. "n" represents the sample size.

indeed available from a random and representative subset of living parasite species. An earlier study identified geographic biases in the genetic knowledge of parasites (Poulin et al., 2019). The present analysis further suggests existing genetic information on helminths is far from random or representative of current biodiversity. Instead, our genetic knowledge of helminth parasites appears biased with respect to which hosts they infect. Factors such as the number of hosts a parasite has, as well as the host's habitat, taxonomic affiliation, and conservation status, significantly influence the likelihood that genetic information has been obtained for that parasite.

Our findings are not really surprising, but they are deeply concerning. On the one hand, it is not surprising that helminth species with many known definitive hosts are more likely to have genetic data (both Nucleotide Sequences and Gene Sequences) in NCBI than those with few hosts. The more host species a parasite has, the more likely it will be the subject of investigation, including genetic studies. Similarly, we should not be surprised that parasites of hosts classified as vulnerable, threatened or endangered have been sequenced disproportionately more frequently than those of hosts of little or no conservation concern: at-risk vertebrate species attract more research attention (Trimble and van Aarde, 2010; Robertson and McKenzie, 2015), some of which must be aimed at their parasites.

On the other hand, our results raise concerns for phylogenetic studies because they suggest biases in the accumulated genetic data on parasites. Far from being a random and representative subset of all species within a large taxonomic group, the available nucleotide sequences used to ascertain the relationships between newly discovered species and previously studied species are determined by the hosts they live in. Representative phylogenetic trees should be based on a representative and unbiased subset of species. Our findings indicate that this is not the case, echoing similar conclusions from studies on vertebrates (Šmíd, 2022; Guedes et al., 2024). Whether genetic data is available for a particular helminth species depends on what host(s) it uses, where they live, and whether their populations are shrinking. The host determines whether parasite sequences are available, not the parasite itself. Existing phylogenies may therefore lack species of great evolutionary significance.

Crucially, biased genetic knowledge of parasites can limit our ability to resolve the evolution of some of their key features. For example, host specificity, i.e. the number of host species a parasite can use at a given stage in its life cycle, is an important trait of parasites determining their likelihood of host-switching as well as their risk of extinction (Poulin et al. 2011). Trends in the evolution of host specificity, for example what factors favor a relaxation of specificity, or whether the evolution of strict host specificity is reversible, have been investigated in well-studied groups or within particular clades with a limited geographic range (see Desdevises et al. 2002; Poulin et al. 2006; Mendlová and Šimkova 2014). This generally requires mapping host specificity along the branch tips (extant species) on a well-resolved parasite phylogenetic tree. However, our results suggest that any attempt to resolve the evolution of host specificity for helminth parasites on larger scales would be flawed, because host-specific parasite species (those with few host species) are underrepresented in the NCBI database. Any parasite phylogeny based on available sequence data would not capture the true distribution of host specificity among living parasite species, and would thus lead to erroneous conclusions about its evolution.

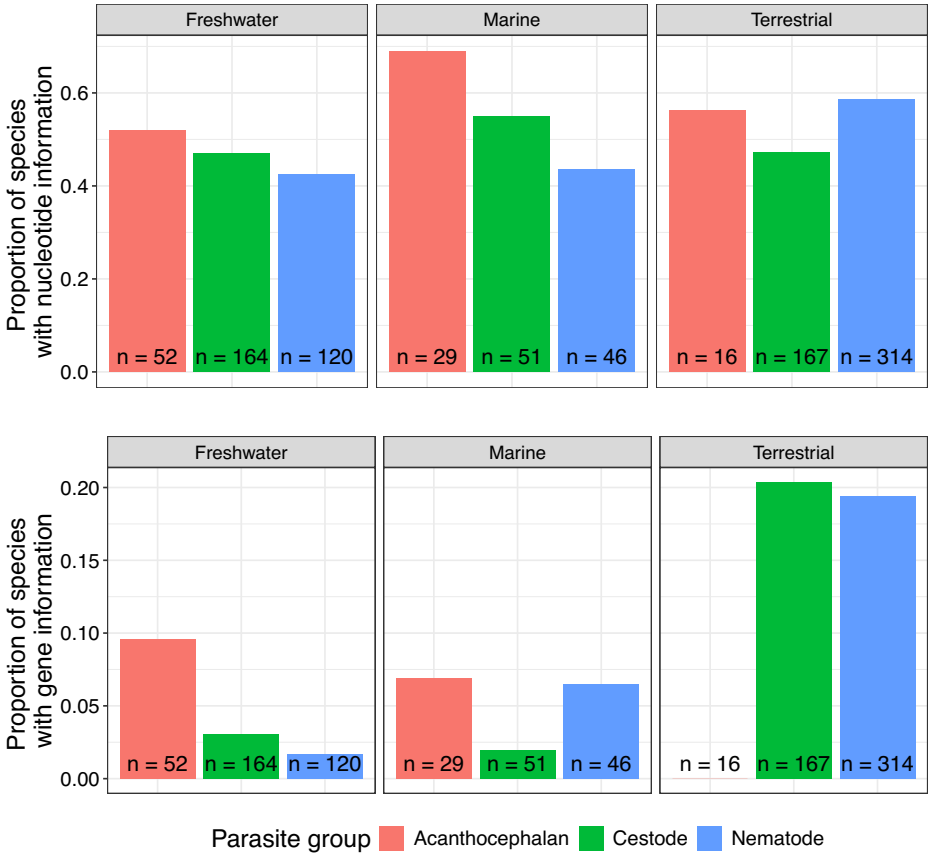

Fig. 4. Genetic data availability by habitat and parasite group. Proportion of helminth parasite species for which genetic information is available either in the NCBI Nucleotide database (top) or Gene database (bottom) as a function of their habitat (freshwater, marine, and terrestrial) and parasite group (acanthocephalan, cestode, and nematode). "n" represents the sample size.

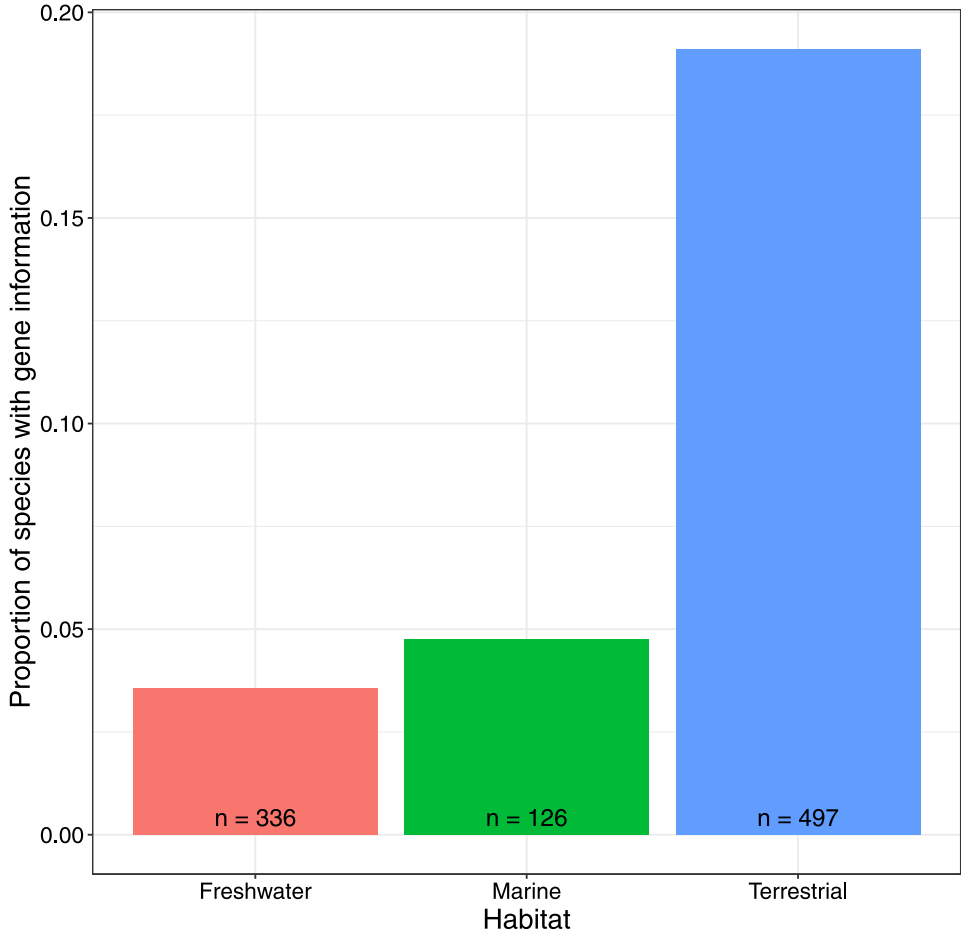

Fig. 5. Genetic data availability by habitat. Proportion of helminth parasite species for which genetic information is available or not in the NCBI Gene database a function of their habitat (freshwater, marine, and terrestrial). "n" represents the sample size.

Incomplete information is a common problem in phylogenetically-based evolutionary analysis. When genetic data are available for a species and it can be included in a phylogenetic tree, but no information is available on a focal trait for that species, the missing trait value can be imputed by extrapolation from the values of closely-related species (Debastiani et al. 2021). Thus, traits can be estimated from a species' phylogenetic position. However, the opposite is not the case: we cannot determine the phylogenetic placement of a species from its physiological or ecological traits. Therefore, the biased availability of genetic data for parasites with respect to the hosts they exploit limits any attempt to explore the evolution of host specificity or other parasite traits related to host use.

We acknowledge that our results are derived from existing databases, and not from data that capture the full inventory of all living parasite species. However, we cannot think of any explanation for why parasites with different types of hosts included in a large database compiled independently (Benesh et al., 2017) would have different likelihood of having genetic data in NCBI, other than the apparent biases we report in the present study. We are therefore confident that our results reveal a true pattern.

In summary, our findings demonstrate that the probability that genetic data are available for parasite species is significantly linked to the hosts they use, and perhaps not so much to the parasite's own properties. This has important implications. First, not only is the placement of many helminth species within phylogenies uncertain (the so-called Darwinian shortfall; Diniz-Filho et al., 2013), but also those with missing genetic information are over-represented among parasites with higher host specificity and/or exploiting hosts of certain taxa or of no conservation concern. Second, the biased availability of genetic information for helminths limits our ability to resolve the evolutionary direction or drivers of several of their key traits within a phylogenetic context. No doubt the coming years will see the continuing growth of genetic databases; however, as we show here, at present our genetic knowledge of parasitic helminths is not truly representative of their biodiversity.

Data availability statement

The full dataset used in this study is available on Figshare (https://figshare.com/s/d39b4988c6c1a54d136c?file=51785765).

CRediT authorship contribution statement

Chen-Hua Li: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Robert Poulin:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Conceptualization.

Acknowledgements

We thank Jerusha Bennett and Priscila Salloum for useful and constructive comments on an earlier version.

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.ijpara.2025.05.007.

References

Benesh, D.P., Lafferty, K.D., Kuris, A., 2017. A life cycle database for parasitic acanthocephalans, cestodes, and nematodes. Ecology 98, 882.

Blasco-Costa, I., Poulin, R., 2013. Host traits explain the genetic structure of parasites: A meta-analysis. Parasitology 140, 1316–1322.

Blasco-Costa, I., Cutmore, S.C., Miller, T.L., Nolan, M.J., 2016. Molecular approaches to trematode systematics: "Best practice" and implications for future study. Syst. Parasitol. 93, 295–306.

Caira, J.N., Jensen, K., Waeschenbach, A., Olson, P.D., Littlewood, D.T., 2014. Orders out of chaos: Molecular phylogenetics reveals the complexity of shark and stingray tapeworm relationships. Int. J. Parasitol. 44, 55–73.

Debastiani, V.J., Bastazini, V.A.G., Pillar, V.D., 2021. Using phylogenetic information to impute missing functional trait values in ecological databases. Ecol. Inform. 63, 101315.

Desdevises, Y., Morand, S., Legendre, P., 2002. Evolution and determinants of host specificity in the genus *Lamellodiscus* (Monogenea). Biol. J. Linn. Soc. 77, 431–443.

Diniz-Filho, J.A.F., Loyola, R.D., Raia, P., Mooers, A.O., Bini, L.M., 2013. Darwinian shortfalls in biodiversity conservation. Trends Ecol. Evol. 28, 689–695.

Guedes, J.J.M., Diniz-Filho, J.A.F., Moura, M.R., 2024. Macroecological correlates of Darwinian shortfalls across terrestrial vertebrates. Biol. Lett. 20, 20240216.

IUCN, 2024. The IUCN Red List of Threatened Species. Version 2024-2. https://www.iucnredlist.org. Accessed on 27 November 2024.

Jetz, W., Pyron, R.A., 2018. The interplay of past diversification and evolutionary isolation with present imperilment across the amphibian tree of life. Nat. Ecol. Evol. 2, 850–858.

Mendlová, M., Šimkova, A., 2014. Evolution of host specificity in monogeneans parasitizing Africal cichlid fish. Parasit. Vect. 7, 69.

Nadler, S.A., Pérez-Ponce de León, G., 2011. Integrating molecular and morphological approaches for characterizing parasite cryptic species: Implications for parasitology. Parasitology 138, 1688–1709.

Olson, P.D., Tkach, V.V., 2005. Advances and trends in the molecular systematics of the parasitic Platyhelminthes. Adv. Parasitol. 60, 165–243.

Pérez-Ponce de León, G., Hernández-Mena, D.I., 2019. Testing the higher-level phylogenetic classification of Digenea (Platyhelminthes, Trematoda) based on nuclear rDNA sequences before entering the age of the 'next–generation' tree of life. J. Helminthol. 93, 260–276.

Pérez-Ponce de León, G., Poulin, R., 2018. An updated look at the uneven distribution of cryptic diversity among parasitic helminths. J. Helminthol. 92, 197–202.

Perkins, S.L., Martinsen, E.S., Falk, B.J., 2011. Do molecules matter more than morphology? Promises and pitfalls in parasites. Parasitology 138, 1664–1674.

Poulin, R., Krasnov, B.R., Shenbrot, G.I., Mouillot, D., Khokhlova, I.S., 2006. Evolution of host specificity in fleas: is it directional and irreversible? Int. J. Parasitol. 36, 185–191.

Poulin, R., Krasnov, B.R., Mouillot, D., 2011. Host specificity in phylogenetic and geographic space. Trends Parasitol. 27, 355–361.

Poulin, R., Hay, E., Jorge, F., 2019. Taxonomic and geographic bias in the genetic study of helminth parasites. Int. J. Parasitol. 49, 429–435.

Poulin, R., Presswell, B., 2016. Taxonomic quality of species descriptions varies over time and with the number of authors, but unevenly among parasitic taxa. Syst. Biol. 65, 1107–1116.

Poulin, R., Presswell, B., Bennett, J., de Angeli Dutra, D., Salloum, P.M., 2023. Biases in parasite biodiversity research: why some helminth species attract more research than others. Int. J. Parasitol. – Parasit Wildl. 21, 89–98.

R Core Team, 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria https://www.R-project.org/.

Robertson, P.A., McKenzie, A.J., 2015. The scientific profiles of terrestrial mammals in Great Britain as measured by publication metrics. Mamm. Rev. 45, 128–132.

Selbach, C., Jorge, F., Dowle, E., Bennett, J., Chai, X., Doherty, J.-F., Eriksson, A., Filion, A., Hay, E., Herbison, R., Lindner, J., Park, E., Presswell, B., Ruehle, B., Sobrinho, P. M., Wainwright, E., Poulin, R., 2019. Parasitological research in the molecular age. Parasitology 146, 1361–1370.

Sherrill-Mix, S., 2024. taxonomizr: Functions to Work with NCBI Accessions and Taxonomy. R package version 0.10.6, https://github.com/sherrillmix/taxonomizr.

Šmíd, J., 2022. Geographic and taxonomic biases in the vertebrate tree of life. J. Biogeogr. 49, 2120–2129.

Thorn, C.S., Maness, R.W., Hulke, J.M., Delmore, K.E., Criscione, C.D., 2023. Population genomics of helminth parasites. J. Helminthol. 97, e29.

Trimble, M.J., van Aarde, R.J., 2010. Species inequality in scientific study. Conserv. Biol. 24, 886–890.

Upham, N.S., Esselstyn, J.A., Jetz, W., 2019. Inferring the mammal tree: species-level sets of phylogenies for questions in ecology, evolution, and conservation. PLoS Biol. 17, e3000494.

Velázquez-Urrieta, Y., Mendoza-Portillo, V., García-De León, F.J., 2024. Diversity of trematodes (Platyhelminthes) in Mexico with an assessment of the availability of genetic data for their conservation. J. Helminthol. 98, e92.

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.