International Journal for Parasitology: Parasites and Wildlife 21 (2023) 89–98

Contents lists available at ScienceDirect

International Journal for Parasitology: Parasites and Wildlife

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

Biases in parasite biodiversity research: why some helminth species attract more research than others

Robert Poulin *, Bronwen Presswell, Jerusha Bennett, Daniela de Angeli Dutra, Priscila M. Salloum

Department of Zoology, University of Otago, P.O. Box 56, Dunedin, 9054, New Zealand

ARTICLE INFO

Keywords:
Conservation
Human uses
Number of authors
Research effort
Species knowledge
Taxonomy

ABSTRACT

As the number of known and described parasite species grows every year, one might ask: how much do we actually know about these species beyond the fact they exist? For free-living taxa, research effort is biased toward a small subset of species based on their properties or human-centric factors. Here, using a large data set on over 2500 helminth parasite species described in the past two decades, we test the importance of several predictors on two measures of research effort: the number of times a species description is cited following its publication, and the number of times a species' name is mentioned in the scientific literature. Our analysis highlights some taxonomic biases: for instance, descriptions of acanthocephalans and nematodes tend to receive more citations than those of other helminths, and species of cestodes are less frequently mentioned in the literature than other helminths. We also found that helminths infecting host species of conservation concern receive less research attention, perhaps because of the constraints associated with research on threatened animals, while those infecting host species of human use receive greater research effort. Intriguingly, we found that species originally described by many co-authors subsequently attract more research effort than those described by one or few authors, and that research effort correlates negatively with the human population size of the country where a species was discovered, but not with its economic strength, measured by its gross domestic product. Overall, our findings reveal that we have conducted very little research, or none at all, on the majority of helminth parasite species following their discovery. The biases in study effort we identify have serious implications for future research into parasite biodiversity and conservation.

1. Introduction

There is general agreement that there remain a huge number of parasite species yet to be discovered and described before we complete our inventory of extant parasite biodiversity (Poulin, 2014; Jorge and Poulin, 2018; Carlson et al., 2020a). Attempts to assess the current state of our knowledge of parasite diversity usually ask how many species are left to be found, or what proportion of total biodiversity on Earth consists of parasites (Dobson et al., 2008; Poulin, 2014; Carlson et al., 2020a). However, in addition to asking such questions, it is equally important to ask how much we actually know about the species we have already found. The emphasis is often on discovering as many species as possible and not on learning more about them following their discovery. In other words, breadth of knowledge is often prioritised over depth of knowledge in biodiversity research. Here, we focus on the latter. After first being discovered, described and assigned a unique Latin name, how often are parasite species seen again, studied again, or even mentioned again in the scientific literature?

Research effort is notoriously allocated unevenly across known species, with species that are 'easy' to study (locally abundant, accessed without difficulty or easily maintained in captivity) receiving the most attention (Westoby, 2002). Other biases also plague biodiversity research. Firstly, properties of the organisms themselves can determine whether they will be the subject of much further research. These include simple higher-level taxonomic effects. For example, among animals, vertebrates are the subject of greater research effort relative to their diversity than invertebrates (Titley et al., 2017). Within higher taxa, species-level properties also matter. Among mammals, the body size, native versus introduced status, conservation status, economic relevance and human uses of individual species combine to determine how much research attention they attract as measured by their mention in the scientific literature (Trimble and van Aarde, 2010; Robertson and

* Corresponding author.
E-mail address: robert.poulin@otago.ac.nz (R. Poulin).

https://doi.org/10.1016/j.ijppaw.2023.04.010
Received 2 April 2023; Received in revised form 26 April 2023; Accepted 26 April 2023
Available online 28 April 2023

McKenzie, 2015; Fleming and Bateman, 2016; dos Santos et al., 2020; Tam et al., 2022). Similar trends emerge from an analysis of research effort on reptiles (Guedes et al., 2023). Research effort can also be driven by purely subjective criteria. For example, flower colour predicts how much research attention is directed toward particular plant species better than more fundamental ecological traits (Adamo et al., 2021).

Factors unrelated to the species themselves can also determine the amount of attention they receive from scientists. For instance, socio-economic factors and authorship of the original species description may also be influential. A disproportionate amount of biodiversity research is aimed at species occurring in temperate countries with larger economies, and conducted by researchers from wealthier nations (Titley et al., 2017; dos Santos et al., 2020), indicating that the scientific capacity of the country where a species is found is a strong driver of research effort. Also, the number of individual researchers involved in the description of a new species may influence how often it will be the focus of future research. Indeed, as a general rule in the ecological literature, the greater the number of co-authors of a particular article, the greater the number of citations it receives in the years after its publication (Leimu and Koricheva, 2005; Borsuk et al., 2009; Fox et al., 2016).

For these and other reasons, in any higher taxon of free-living organisms, not all species are equal: a small number of species command most of the scientific attention. Is this inequality in research effort and accumulated knowledge also applicable to parasites? And what kinds of parasites are subject to the most research following their discovery? Further, in the case of parasites, host characteristics may also determine research effort. Here, we use bibliometric data as measures of research effort on particular parasite species, to assess how much research attention they have received, and what factors explain variation in research effort among parasite species. Specifically, we test the influence of parasite and host taxonomy, authorship of the species description, host properties (body size, conservation status, human uses), and socioeconomic factors (gross domestic product and population size of the country of species discovery) on the subsequent research effort directed at helminth parasites discovered and described since the year 2000. Our findings reveal that most parasite species receive little or no attention at all following their discovery, and also identify the key factors that determine which privileged species do in fact attract further research.

2. Materials and methods

2.1. Dataset compilation

As the basis for our dataset, we used all species of helminth parasites described between 2000 and 2018 inclusively in the following 8 journals: *Acta Parasitologica* (data from 2000 to 2005 missing for this journal), *Comparative Parasitology*, *Folia Parasitologica*, *Journal of Helminthology*, *Journal of Parasitology*, *Parasitology International*, *Parasitology Research*, and *Systematic Parasitology*. This was extracted from the more extensive compilations in Poulin et al. (2022a, 2022b). We stopped at 2018 to allow time post-description for additional research to be conducted on a species; very recently discovered species are unlikely to have been the subject of any further research. For each species, we recorded the higher taxon to which it belonged (acanthocephalans, monogeneans, cestodes, trematodes, or nematodes), the Latin name of its type-host species, the higher taxon to which the host belonged (invertebrates, fish [including elasmobranchs], amphibians, reptiles, birds, or mammals), the country where it was first discovered (based on the type locality), the number of authors of the published species description, and the year and journal in which it was published. In a few cases, the parasite was found as a larval stage in an intermediate host and its adult form was obtained through experimental infection of a lab animal (e.g., rat, chicken); since the natural definitive host was unknown, these species were excluded from the dataset. Also, in a few cases, the species were obtained from fish caught in international waters; in these cases, the country of origin was assigned as that where the authors were based.

We then obtained two different metrics of research effort for each parasite species from the Zoological Record™ database, searched in January 2023. First, we recorded the number of times the article presenting the original species description was cited, the minimum being zero. Second, we conducted a topic search using the Latin binomial name of the species (genus and species names combined, not separate), and recorded the number of hits as representing the number of times the species was mentioned in the scientific literature following its description. The minimum value was also zero, since for this measure we excluded the paper in which it was originally described. These two measures of research effort are positively correlated across all species in our dataset (Spearman's rank correlation, $r_s = 0.2578$, $N = 2536$, $P < 0.0001$), but only weakly, therefore we used them as independent measures in separate analyses.

The search that produced the second measure above yielded many articles in which the species was mentioned without new specimens necessarily being studied; perhaps the species was only mentioned in the context of a taxonomic or phylogenetic study of related species. We validated this measure of research effort using a subset of species: those described in the arbitrarily chosen years 2006-2007. For these species, we carefully read the abstract, and in some cases the article itself, of all articles in which the name of the species was mentioned, to determine in how many of those (the minimum was zero) new specimens of the focal species had actually been obtained and/or studied as part of biodiversity surveys, genetic studies, laboratory experiments, etc. This provides a more conservative and accurate estimate of how much genuine new knowledge has been obtained about a species. We found a strong relationship between this latter number and the number of articles in which a species is simply mentioned based on the topic search described above (Spearman's rank correlation, $r_s = 0.789$, $N = 329$, $P < 0.001$). Based on this relatively strong correlation coefficient, the number of articles returned by the search provides a good proxy for the amount of research conducted on a species following its discovery and description. One minor caveat is that in the years following their description, some helminth species may have been re-assigned to a different genus, or synonymised with a previously described species, and by searching only publications using their original Latin name, we may have underestimated how often they were the focus of research. However, this is likely to apply to a very small proportion of the large number of species considered here.

In addition to the (i) parasite higher taxon, (ii) the host higher taxon, (iii) the number of authors of the species description and (iv) the number of years since it was published, we also obtained data on further predictors of research effort (see Supplementary Material for full details). Since, all else being equal, larger animals generally receive more research attention, this may indirectly extend to their parasites, therefore (v) host body mass was obtained by matching the Latin name of the type-host species with corresponding data from various databases and publications. In some cases, only the genus of the type-host was given, or the type-host was named but no body mass information was available for that species in the databases we searched; in both these situations, we used the average body mass of all species in the host genus or family as the best estimate of the body mass for the type-host. For some fish and invertebrate species, body mass had to be estimated from published mass-versus-length regression equations (see Supplementary Material). Because the conservation status and human uses of host species also affect how much research they command, (vi) the IUCN (International Union for Conservation of Nature) Red List category and (vii) Human Use for each type-host species were obtained from the IUCN database (https://www.iucnredlist.org/). For analysis, we considered three categories for conservation status: Species of conservation concern (lumping together those in the IUCN categories Near threatened, Vulnerable, Conservation dependent, Endangered, and Critically endangered), Species of least concern, and Species that are data deficient (including those not listed in the IUCN database). Similarly, for analysis we considered two categories regarding human use: Species of no recognised human

Table 1
Number of parasite species included in our analysis, broken down by higher parasite taxon and host taxon.

	Trematodes	Cestodes	Monogeneans	Nematodes	Acanthocephalans	TOTAL
Invertebrates	2	0	0	54	0	56
Fish	422	332	588	284	69	1695
Amphibians	20	3	14	81	7	125
Reptiles	41	18	6	148	6	219
Birds	66	33	0	38	20	157
Mammals	25	66	0	186	7	284
TOTAL	576	452	608	791	109	2536

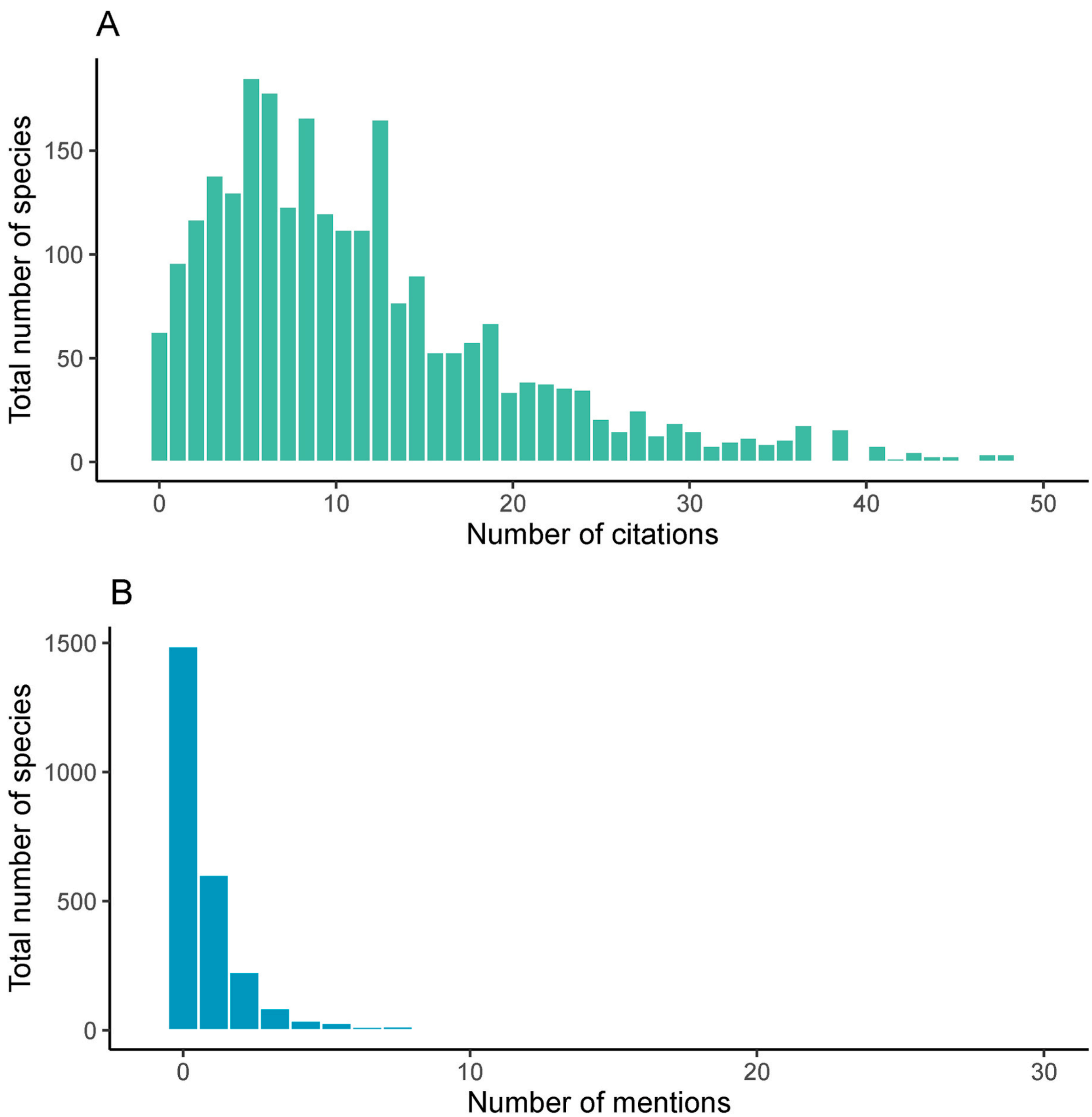

Fig. 1. Frequency distribution of (A) number of citations received by species descriptions following their publication, and (B) number of mentions of a species' name in the scientific literature following its description, for helminth parasites described between the years 2000 and 2018. Data from the Zoological Record™ database. Extreme values (25 species descriptions with more than 50 citations; 3 species with more than 30 mentions) are excluded to avoid distorting the figures.

use, and Species of human use (lumping together all IUCN End Use categories, i.e. food, fibre, handicrafts and jewellery, pets, research, sport hunting, etc.). Finally, to account for the influence of socioeconomic factors on research capacity, we also obtained the (viii) gross domestic product (GDP) averaged across the years 2000 and 2018, and (ix) human population size, also averaged across the years 2000 and 2018, for the country in which each parasite species was first discovered, obtained using the R package WDI (https://github.com/vincentarelbundock/WDI) which extracts information from databases hosted by the World Bank (see Supplementary Material). Some studies on inequality in research effort into animal species have also identified latitude as influential, however latitude is positively correlated with GDP

Table 2
Results of a generalized linear mixed model testing the effects of various predictors on the number of citations of the original species description in Zoological Record™. For categorical predictors with more than two levels, the reference level was chosen arbitrarily (parasite taxon = acanthocephalans, host taxon = amphibians, conservation status = of least concern); selecting a different reference level had little impact on the results. The journal in which a species description was published (random factor) accounted for 2% of unexplained variance. Significant effects (NB: based on uncorrected *P*-values) are shown in bold.

Predictor	Estimate	Standard error	z-value	*P*
Intercept	**2.589**	**0.114**	**22.725**	**<0.0001**
Parasite taxon: Monogeneans	−0.152	0.081	1.870	0.0615
Parasite taxon: Cestodes	**−0.216**	**0.083**	**2.598**	**0.0094**
Parasite taxon: Trematodes	**−0.210**	**0.080**	**2.618**	**0.0089**
Parasite taxon: Nematodes	−0.058	0.079	0.729	0.4659
Host taxon: Invertebrates	**−0.550**	**0.129**	**4.258**	**<0.0001**
Host taxon: Fish	**0.190**	**0.075**	**2.540**	**0.0111**
Host taxon: Reptiles	−0.047	0.086	0.547	0.584
Host taxon: Birds	**−0.247**	**0.096**	**2.563**	**0.0104**
Host taxon: Mammals	**−0.169**	**0.084**	**2.006**	**0.0449**
Number of authors of the species description	**0.131**	**0.016**	**8.031**	**<0.0001**
Number of years since the description was published	**0.274**	**0.018**	**15.684**	**<0.0001**
Host body mass	0.022	0.017	1.329	0.1839
Conservation status: Data deficient	0.059	0.043	1.375	0.1692
Conservation status: Of concern	**−0.178**	**0.048**	**3.702**	**0.0002**
Human use: Of no recognised human use	**−0.106**	**0.038**	**2.778**	**0.0055**
Country's GDP	0.007	0.017	0.434	0.6643
Country's human population size	**−0.152**	**0.017**	**8.771**	**<0.0001**

(Nordhaus, 2006), and the latter is a more direct measure of a country's research capacity.

2.2. Data analysis

We tested the effect of predictors of research effort using generalized linear mixed models (GLMMs), using the *glmmTMB* package (https://github.com/glmmTMB/glmmTMB) in the R computing environment (R Core Team, 2022). For each of our two response variables (number of citations of the species description, and number of mentions of the species' name post-description), data were modelled with a negative binomial distribution, to account for the large number of zero values. We used the same 9 predictors (fixed factors) in both GLMMs: (i) parasite taxon (five levels: acanthocephalans, monogeneans, cestodes, trematodes, or nematodes); (ii) host taxon (six levels: invertebrates, fish, amphibians, reptiles, birds, or mammals); (iii) number of authors of the description; (iv) number of years since the description was published; (v) host body mass; (vi) conservation status (three levels: of conservation concern, of least concern, data deficient); (vii) human use category (two levels: of human use, of no recognised human use); (viii) GDP of the country where it was found; (ix) human population size of the country where it was found. Both GLMMs included the journal as a random factor, to account for any variation among journals in how likely the research they publish gets noticed and cited in subsequent years.

3. Results

Our dataset consisted of data on 2536 helminth species described between the years 2000 and 2018 inclusively. The majority were nematodes, followed by monogeneans and trematodes, whereas the most common host species by far were fish (Table 1). These helminth species were described in papers with number of authors ranging from 1 up to 13. Overall, most host species (1451; 57.2%) of helminths in our dataset are classified as of least concern by the IUCN, 387 (15.3%) are listed as of concern, and the rest (698; 27.5%) are data deficient. From the perspective of their use to humans, 1234 (48.7%) host species are listed as having some use to humans, while the rest (1302; 51.3%) have no recognised human use (see Supplementary Material for full dataset).

Our two measures of research effort are both strongly skewed toward small values (Fig. 1). Although papers describing helminths have not been cited for only 63 (2.5%) out of 2536 species in our dataset, a shocking 1488 species (58.7%) have not been mentioned again in the scientific literature following their original description. At the other end of the spectrum (i.e., maximum values), the paper describing the nematode *Anisakis berlandi*, published in 2014, has been cited 138 times (as of January 2023), while the trematode *Maritrema novaezealandensis*, described in the year 2004, has been mentioned in 51 other publications since its description.

Several of the predictors tested in the GLMM were related to how often a paper describing a species was cited (Table 2). The GLMM and the results of pairwise comparison (Tukey tests, not shown) indicate that there were some differences among taxa; in particular, papers describing acanthocephalans and nematodes tended to receive more citations than those describing other types of helminths (Fig. 2A), and papers describing helminths of fish hosts were cited more than those describing species from other host groups, especially invertebrates (Fig. 3A). Not surprisingly, the number of years since a species description was published was strongly and positively related to how many citations it received. The number of authors of the original species description was also positively correlated with how often it was subsequently cited (Fig. 4A). With respect to host properties, the body mass of the type-host species did not affect the number of citations received, however descriptions of parasites from host species of conservation concern received fewer citations than those for hosts of different conservation status, and those of hosts with no recognised human use also received fewer citations than those for hosts with some human use (Fig. 5). Finally, the number of times a species description was cited was not influenced by the GDP of the country where it was found, but it correlated negatively with the country's population size.

In the second GLMM, using as response variable the number of mentions of a species in the scientific literature following publication of its original description, fewer predictors emerged as important (Table 3). Again, based on the GLMM and the results of pairwise comparison (Tukey tests, not shown), the most notable taxonomic differences are that cestodes tend to receive fewer mentions after their descriptions than other helminth taxa (Fig. 2B), and that helminths from fish hosts also tend to have fewer mentions in the literature than helminths from other host groups (Fig. 3B). Again, both the number of years since a species description was published and the number of authors of the original description were positively (only weakly, in the latter case) related to how many times the species was subsequently mentioned in the scientific literature (Fig. 4B). Finally, none of the host properties considered (body mass, conservation status, and human use status) or the properties of the country of discovery (GDP, population size) had any effect on the number of times a species was mentioned in the literature following its description.

4. Discussion

The focus of parasite biodiversity research on either achieving a full inventory of extant species or just estimating how many there are on Earth is commendable, but logistically challenging (Poulin, 2014; Carlson et al., 2020a) and perhaps even unattainable (Stropp et al., 2022). In contrast, obtaining further information on the species we already know can shed light on multiple aspects of parasite biology, from their evolutionary history to their ecological impacts. There is currently growing interest in parasite conservation (Gómez and Nichols, 2013; Carlson et al., 2020b), however, to conserve parasite species, we must first know about them. Here, we demonstrate that when they are

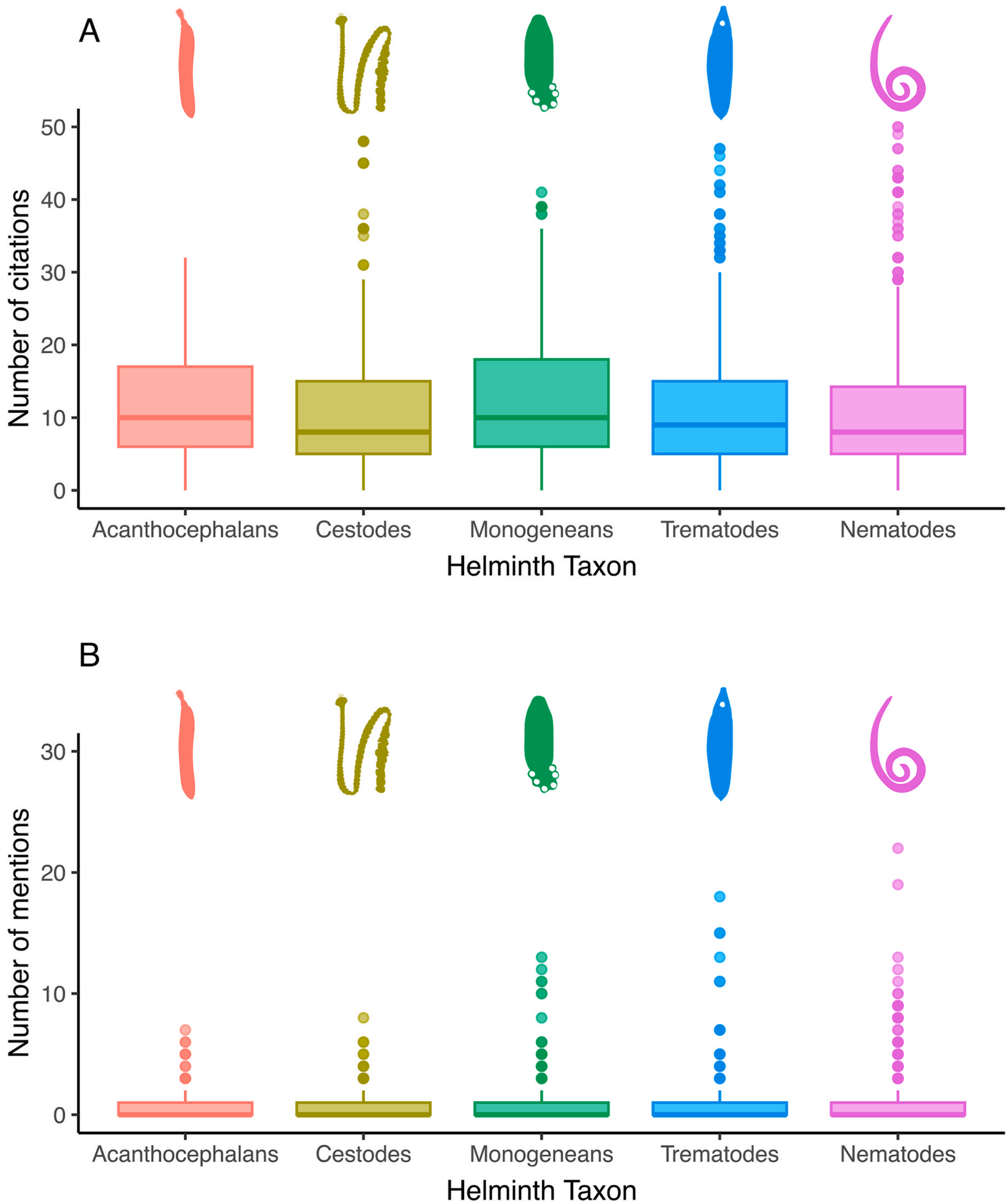

Fig. 2. Box plots (median and interquartile range) showing (A) number of citations received by species descriptions following their publication, and (B) number of mentions of a species' name in the scientific literature following its description, for helminth parasites of five higher taxa. Data from the Zoological Record™ database. Values greater than 50 citations (comprising 1 cestode, 10 monogeneans, 3 trematodes, and 11 nematodes) and greater than 30 mentions (comprising 1 trematode and 2 nematodes) are excluded to avoid distorting the figures.

first discovered and described, parasite species do not enter an even playing field: most will be totally ignored or receive only limited attention in the years following their description, whereas a few will be the subject of much additional research.

The lack of research on most helminth species is striking. Nearly 60% of the species in our dataset have not been mentioned again in the scientific literature following their description. Further evidence of the lack of parasite species knowledge accruing over time after their description is that, for the 329 species described between 2006 and 2007, 72% (237) of them had no new specimen collected and studied in the subsequent 15 years, according to the Zoological Record™ database. Thus, our search revealed that after it has been described, a typical parasite species is not studied again for a long time, if ever.

The two metrics of research effort we used are only weakly correlated (Spearman $rs \approx 0.25$), and thus capture slightly different aspects of ongoing research. On the one hand, the published description of a species may receive many citations without the species' name being mentioned. For instance, the paper may be cited as a past example of

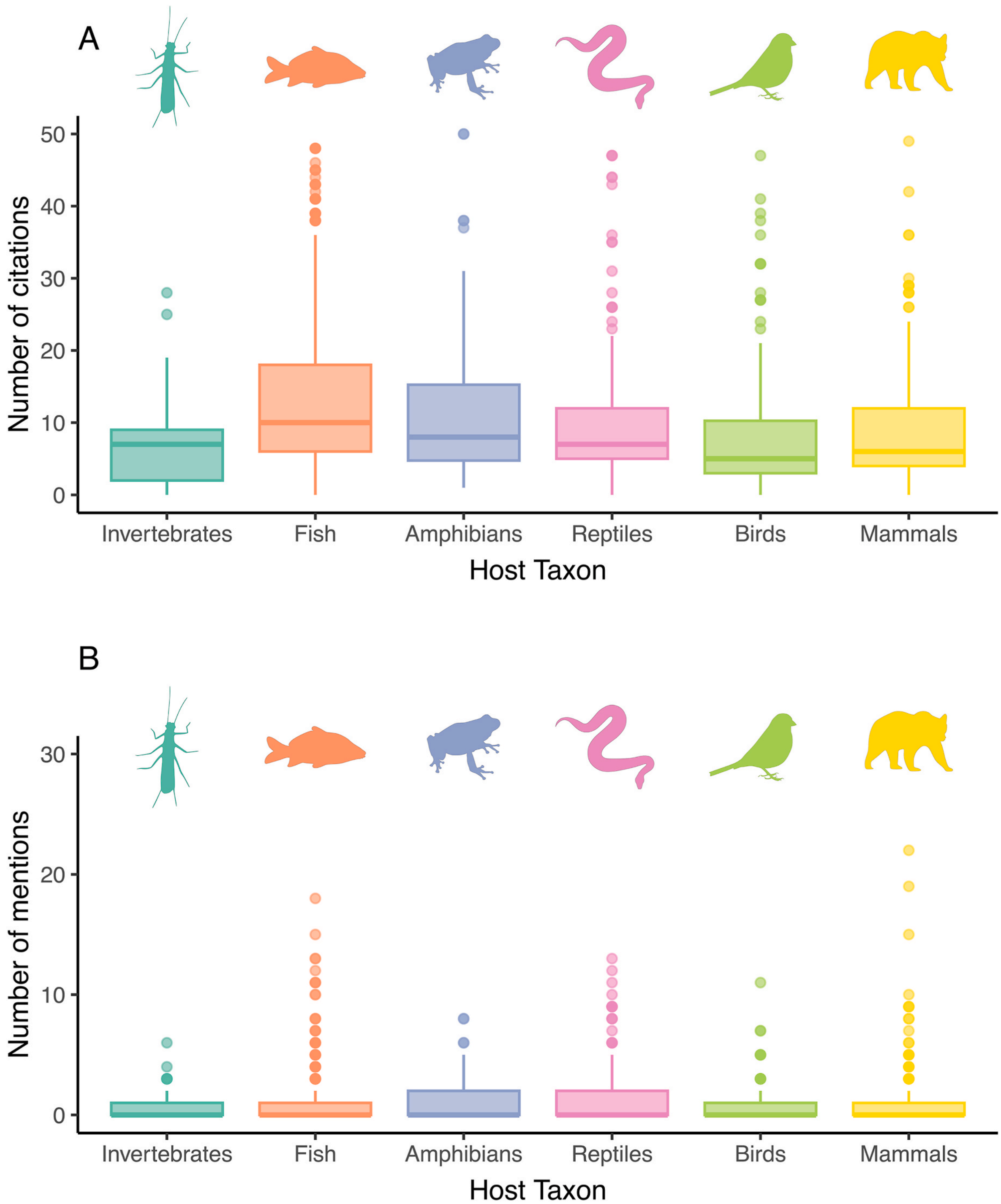

Fig. 3. Box plots (median and interquartile range) showing (A) number of citations received by species descriptions following their publication, and (B) number of mentions of a species' name in the scientific literature following its description, for helminths parasitising six host groups. Data from the Zoological Record™ database. Values greater than 50 citations (comprising 14 fish, 1 amphibian, 5 reptiles, 1 bird and 4 mammals) and greater than 30 mentions (comprising 1 amphibian, 1 bird and 1 mammal) are excluded to avoid distorting the figures.

research on that parasite's family, or on its type-host species. On the other hand, a species may receive multiple mentions in the scientific literature without its original description being cited each time. Because the number of mentions correlates rather strongly (based on the 2006-2007 subset of 329 species: Spearman $rs \approx 0.8$) with the number of studies in which new specimens of the species are actually obtained and used in research, the number of mentions of a species' name in the literature is another good measure of the research attention it attracts. It is not unusual for researchers not to refer to the original description of the species whose biology they are investigating.

Not surprisingly, we found that the earlier a species has been described, the more citations the original description has received, and the more mentions its name has received in the literature. More intriguingly, the number of authors of the original species description had a clear positive effect on the number of times that original paper was cited, and a weaker but significant positive effect on the number of times the species' name was subsequently mentioned in the literature. This finding parallels a general trend in ecology, where the greater the

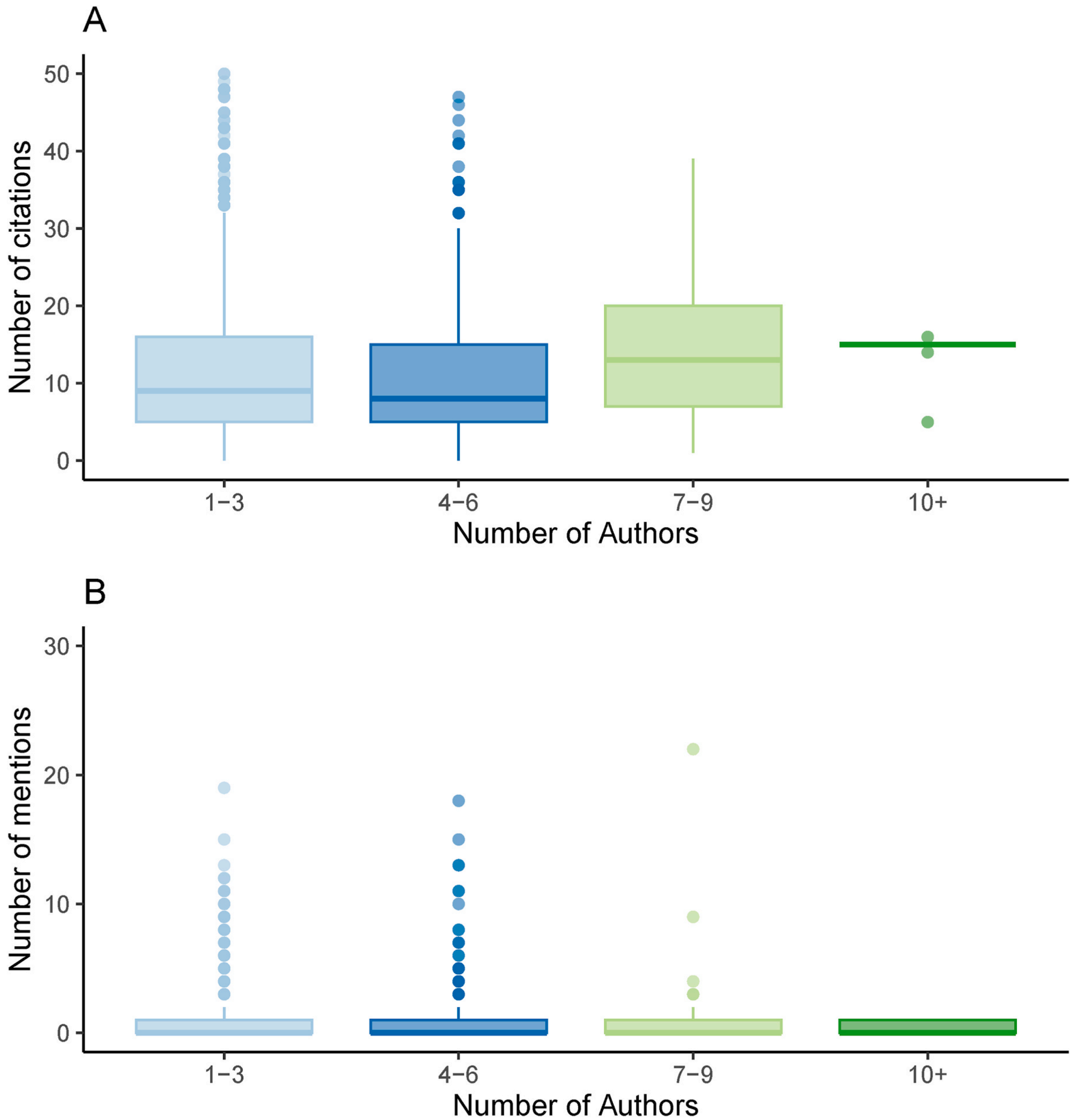

Fig. 4. Box plots (median and interquartile range) showing (A) number of citations received by species descriptions following their publication, and (B) number of mentions of a species' name in the scientific literature following its description, for helminth parasites as a function of the number of authors of the original species description. Data from the Zoological Record™ database. Extreme values (25 species descriptions with more than 50 citations; 3 species with more than 30 mentions) are excluded to avoid distorting the figures.

number of co-authors of a particular article, the more citations it receives in the years after its publication (Leimu and Koricheva, 2005; Borsuk et al., 2009; Fox et al., 2016). Incidentally, the number of authors per scientific paper has been increasing gradually over time in all branches of science (Wuchty et al., 2007), including in parasite taxonomy (Poulin and Presswell, 2016). However, the link between the number of authors and subsequent citations and species mentions suggests that it is not the importance of a species (from any perspective, e.g., ecological, veterinary, etc.) that drives research on it, but how many people happened to be involved in its description. This can lead to biases in research direction that have no biological justification. For instance, a species of little ecological or conservation significance may receive greater research effort than a species from the same higher taxon and found in the same region that causes substantial pathology and can impact threatened host species, simply because more people were involved in the discovery of the former than the latter. If study effort following species discovery was unrelated to the number of authors of species descriptions, the bias would be less likely and the species of greater immediate concern would be more likely to receive attention.

Our results also reveal that not all higher taxa attract the same research effort. Descriptions of acanthocephalans and nematodes tend to receive more citations than description of other helminths, whereas the names of cestodes are mentioned less frequently in the literature after their description than those of other helminths. An analysis of the literature published two decades ago indicated that there were more publications on acanthocephalans and nematodes, relative to their estimated species diversity, than on other groups of helminths (Poulin, 2002). The present results seem to confirm a slight bias in research effort toward acanthocephalans and nematodes relative to other helminths, the former possibly due to a disproportionate taxonomic effort on this low-diversity taxon, and the latter most likely because of their general pathogenicity and potential as zoonotic agents.

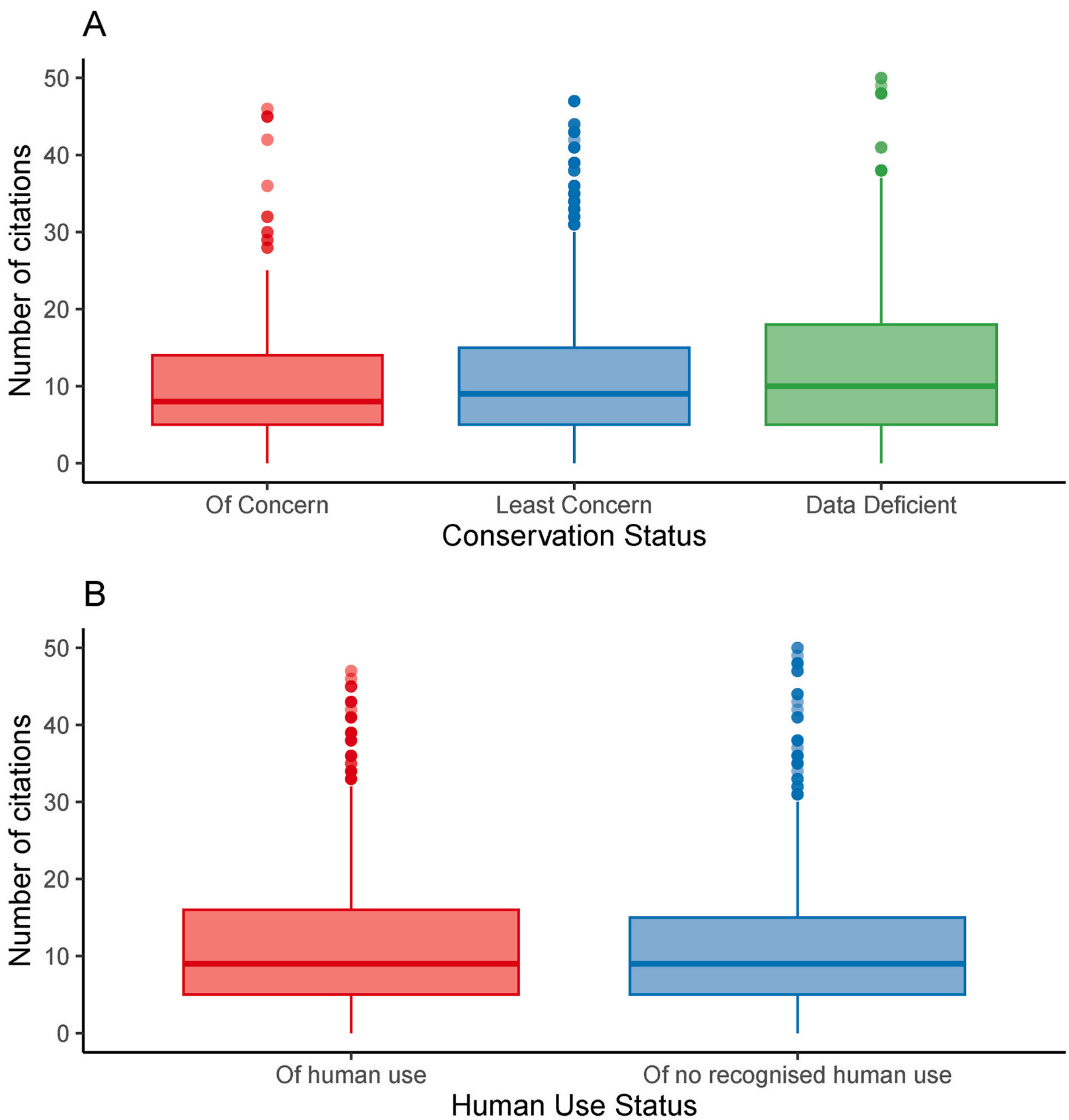

Fig. 5. Box plots (median and interquartile range) showing the number of citations received by species descriptions following their publication as a function of (A) their IUCN conservation status and (B) their IUCN human use status. Data from the Zoological Record™ database. Extreme values (25 species descriptions with more than 50 citations) are excluded to avoid distorting the figures.

From the host perspective, descriptions of helminths recovered from fish hosts were more likely to receive subsequent citations, but in contrast the names of helminths recovered from fish hosts were less likely to appear in the literature, than for other host groups. Given that taxonomic research on fish parasites greatly exceeds that aimed at other parasites (Poulin and Presswell, 2016; Poulin et al., 2020), it is likely that descriptions of fish parasites get regularly cited in subsequent taxonomic papers when previous relevant studies are referenced, without the original species themselves being mentioned, let alone studied again. Not surprisingly, parasites described from invertebrate hosts (in our dataset, always nematodes using insects or snails as hosts) received little further attention relative to those described from vertebrates.

In terms of species-level host properties, the body mass of the type-host did not correlate with the amount of future research on its parasite. This is somewhat surprising, as larger animals generally attract more research attention than small ones (e.g., dos Santos et al., 2020; Tam et al., 2022; Guedes et al., 2023). Also somewhat surprising, we found that species descriptions of helminths from hosts listed as of conservation concern by the IUCN have received fewer citations than those from hosts of no concern or for which there is insufficient data. Some earlier studies have found that vertebrate species of conservation concern tend to receive greater research effort (Trimble and van Aarde, 2010; Robertson and McKenzie, 2015), although the opposite has also been reported (Tam et al., 2022). Regardless, research on parasites of rare or endangered species should be a top priority. However, it may be limited by the stringent regulations associated with threatened or endangered species (Shaw et al., 2021), since this research may require destructive sampling or experimental infection. Finally, we found that species descriptions of helminths from hosts with no recognised human use received fewer citations than those from hosts with human uses. Research on vertebrate hosts themselves is biased toward species with human uses (Tam et al., 2022). There has certainly been extensive research on helminth species described decades before the time span

Table 3
Results of a generalized linear mixed model testing the effects of various predictors on the number of mentions of a species' name in Zoological Record™ following publication of its original description. For categorical predictors with more than two levels, the reference level was chosen arbitrarily (parasite taxon = acanthocephalans, host taxon = amphibians, conservation status = of least concern); selecting a different reference level had little impact on the results. The journal in which a species description was published (random factor) accounted for 5% of unexplained variance. Significant effects (NB: based on uncorrected *P*-values) are shown in bold.

Predictor	Estimate	Standard error	z-value	*P*
Intercept	0.231	0.238	0.969	0.3328
Parasite taxon: Monogeneans	−0.151	0.179	0.846	0.3975
Parasite taxon: Cestodes	**−0.603**	**0.189**	**3.195**	**0.0014**
Parasite taxon: Trematodes	−0.107	0.177	0.606	0.5445
Parasite taxon: Nematodes	−0.094	0.178	0.528	0.5978
Host taxon: Invertebrates	−0.497	0.280	1.774	0.0761
Host taxon: Fish	**−0.429**	**0.157**	**2.734**	**0.0063**
Host taxon: Reptiles	0.042	0.176	0.240	0.8103
Host taxon: Birds	−0.140	0.202	0.691	0.4896
Host taxon: Mammals	0.027	0.173	0.157	0.8756
Number of authors of the species description	**0.215**	**0.037**	**5.872**	**<0.0001**
Number of years since the description was published	**0.486**	**0.038**	**12.798**	**<0.0001**
Host body mass	−0.009	0.042	0.215	0.8297
Conservation status: Data deficient	−0.094	0.095	0.989	0.3225
Conservation status: Of concern	0.067	0.108	0.625	0.5321
Human use: Of no recognised human use	0.062	0.085	0.724	0.4690
Country's GDP	−0.010	0.038	0.270	0.7870
Country's human population size	−0.025	0.038	0.666	0.5052

covered by our dataset that parasitise livestock (e.g., *Haemonchus contortus*; Emery et al., 2016) or pets (*Toxocara* spp.; Overgaauw and van Knapen, 2013) and on species with other impacts on human activities. Our results show that even among more recently described helminth species, whether or not the host is relevant to humans determines the research attention a parasite will receive.

We focused on host properties (body mass, conservation status, human uses) as potential determinants of how much research effort has been directed at their parasites. Certain properties of parasite species themselves could also make them more or less likely to receive further study. For instance, their virulence, their prevalence or their distribution across the geographical range of their host(s) might determine how frequently they are encountered and studied. However, species-level data on such variables are lacking for practically all species in our dataset. Nevertheless, we can still explain why particular species have been well-researched based on their properties. For instance, the species description that was the most highly-cited in our dataset was that of *Anisakis berlandi* (Mattiucci et al., 2014), a potentially zoonotic nematode of concern for commercial fisheries. The parasite whose name has been mentioned the most frequently following its description is the trematode *Maritrema novaezealandensis*; because it can readily be used in laboratory studies, this species has been adopted as an important model species for eco-evolutionary research on topics including population genetics (Keeney et al., 2007), the evolution of host specificity (Koehler et al., 2012), and the impact of global climate change on host-parasite interactions (e.g., Studer et al., 2010; Harland et al., 2015). Several other species ranking highly for one or both of our metrics of study effort belong to taxa known to be pathogenic (e.g., *Anisakis*, *Trichinella*, *Gyrodactylus*). However, such information is lacking for the vast majority of helminth species.

Somewhat surprisingly, parasite species descriptions originating from countries with large populations tended to receive fewer citations than those from smaller countries, with no influence of the country's GDP on citation rates. Both GDP and population size are generally significant predictors of a country's scientific output, each with a positive effect (e.g., Mueller, 2016), thus our results depart from those we expected. This may in part be explained by the fact that our dataset shows the same bias as its source, i.e., the Zoological Record™ database, in that it does not cover many journals that publish papers in languages other than English. It may also be due to the fact that in several cases, the researchers describing a species are based in a different country than that where the parasite was found, creating a possible disconnect between future efforts directed toward that species and the country's population or GDP.

Overall, our findings demonstrate that despite the inventory of known parasite species becoming steadily longer every year (see Poulin, 2014; Poulin and Presswell, 2016), our knowledge of what the species on that growing list actually do is very limited. For the vast majority of these species, we do not know their full life cycle, pathological effects, geographical range, host specificity, etc. This knowledge gulf cannot easily be remedied; it would take vast sums of money and huge investments of time and efforts to acquire basic biological data on even half of the species in our dataset. Therefore, our findings suggest that some of the basic goals of parasite conservation initiatives, which are focused on data collection, risk assessment and prioritisation (Carlson et al., 2020b), may be even more challenging than previously thought.

Funding

This research was not funded by any specific grant from agencies in the public, commercial or not-for-profit sectors.

Author contributions

RP conceived the study, with input from the other authors. RP, BP and PMS compiled the dataset. DdAD conducted the analyses, and DdAD and JB prepared the figures. RP wrote the manuscript with input from all authors. All authors approved the final version and submission of the manuscript.

Declaration of competing interest

The authors declare having no conflict of interest.

Acknowledgements

We are grateful to Alan Lymbery and Andrew Thompson for inviting us to contribute to the special journal issue on 'Parasite Conservation'. We also thank two anonymous reviewers for comments on an earlier version. During the study, DdAD was supported by a publishing bursary from the University of Otago, PMS was supported by funding from the Marsden Fund (Royal Society of New Zealand), and JB was supported through the Cawthron Institute's Ministry of Business, Innovation and Employment, *Emerging Aquatic Diseases* Endeavour grant, award number CAWX2207.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijppaw.2023.04.010.

References

Adamo, M., Chialva, M., Calevo, J., Bertoni, F., Dixon, K., Mammola, S., 2021. Plant scientists' research attention is skewed towards colourful, conspicuous and broadly distributed flowers. Native Plants 7, 574–578.
Borsuk, R.M., Budden, A.E., Leimu, R., Aarssen, L.W., Lortie, C.J., 2009. The influence of author gender, national language and number of authors on citation rate in ecology. Open Ecol. J. 2, 25–28.

Carlson, C.J., Dallas, T.A., Alexander, L.W., Phelan, A.L., Phillips, A.J., 2020a. What would it take to describe the global diversity of parasites? Proc. R. Soc. B 287, 20201841.

Carlson, C.J., Hopkins, S., Bell, K.C., Doña, J., Godfrey, S.S., Kwak, M.L., Lafferty, K.D., Moir, M.L., Speer, K.A., Strona, G., Torchin, M., Wood, C.L., 2020b. A global parasite conservation plan. Biol. Conserv. 250, 108596.

Dobson, A., Lafferty, K.D., Kuris, A.M., Hechinger, R.F., Jetz, W., 2008. Homage to Linnaeus: how many parasites? How many hosts? Proc. Nat. Acad. Sci. USA 105, 11482–11489.

Dos Santos, J.W., Correia, R.A., Malhado, A.C.M., Campos-Silva, J.V., Teles, D., Jepson, P., Ladle, R.J., 2020. Drivers of taxonomic bias in conservation research: a global analysis of terrestrial mammals. Anim. Conserv. 23, 679–688.

Emery, D.L., Hunt, P.W., Le Jambre, L.F., 2016. *Haemonchus contortus*: the then and now, and where to from here? Int. J. Parasitol. 46, 755–769.

Fleming, P.A., Bateman, P.W., 2016. The good, the bad, and the ugly: which Australian terrestrial mammal species attract most research? Mamm Rev. 46, 241–254.

Fox, C.W., Paine, C.E.T., Sauterey, B., 2016. Citations increase with manuscript length, author number, and references cited in ecology journals. Ecol. Evol. 6, 7717–7726.

Gómez, A., Nichols, E., 2013. Neglected wild life: parasitic biodiversity as a conservation target. Int. J. Parasitol. Parasites Wildl. 2, 222–227.

Guedes, J.J.M., Moura, M.R., Diniz-Filho, J.A.F., 2023. Species out of sight: elucidating the determinants of research effort in global reptiles. Ecography 2023, e06491.

Harland, H., MacLeod, C.D., Poulin, R., 2015. Non-linear effects of ocean acidification on the transmission of a marine intertidal parasite. Mar. Ecol. Prog. Ser. 536, 55–64.

Jorge, F., Poulin, R., 2018. Poor geographical match between the distributions of host diversity and parasite discovery effort. Proc. R. Soc. B 285, 20180072.

Keeney, D.B., Waters, J.M., Poulin, R., 2007. Clonal diversity of the marine trematode *Maritrema novaezealandensis* within intermediate hosts: the molecular ecology of parasite life cycles. Mol. Ecol. 16, 431–439.

Koehler, A.V., Springer, Y.P., Randhawa, H.S., Leung, T.L.F., Keeney, D.B., Poulin, R., 2012. Genetic and phenotypic influences on clone-level success and host specialization in a generalist parasite. J. Evol. Biol. 25, 66–79.

Leimu, R., Koricheva, J., 2005. What determines the citation frequency of ecological papers? Trends Ecol. Evol. 20, 28–32.

Mattiucci, S., Cipriani, P., Webb, S.C., Paoletti, M., Marcer, F., Bellisario, B., Gibson, D.I., Nascetti, G., 2014. Genetic and morphological approaches distinguish the three sibling species of the *Anisakis simplex* species complex, with a species designation as *Anisakis berlandi* n. sp. for *A. simplex* sp. C (Nematoda: anisakidae). J. Parasitol. 100, 199–214.

Mueller, C.E., 2016. Accurate forecast of countries' research output by macro-level indicators. Scientometrics 109, 1307–1328.

Nordhaus, W.D., 2006. Geography and macroeconomics: new data and new findings. Proc. Nat. Acad. Sci. USA 103, 3510–3517.

Overgaauw, P.A.M., van Knapen, F., 2013. Veterinary and public health aspects of *Toxocara* spp. Vet. Parasitol. 193, 398–403.

Poulin, R., 2002. Qualitative and quantitative aspects of recent research on helminth parasites. J. Helminthol. 76, 373–376.

Poulin, R., 2014. Parasite biodiversity revisited: frontiers and constraints. Int. J. Parasitol. 44, 581–589.

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., Jorge, F., 2020. The fate of fish parasite discovery and taxonomy: a critical assessment and a look forward. Int. J. Parasitol. 50, 733–742.

Poulin, R., de Angeli Dutra, D., Presswell, B., 2022a. Short and sweet : an analysis of the length of parasite species names. Syst. Parasitol. 99, 699–706.

Poulin, R., McDougall, C., Presswell, B., 2022b. What's in a name? Taxonomic and gender biases in the etymology of new species names. Proc. R. Soc. B 289, 20212708.

R Core Team, 2022. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

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.

Shaw, R.C., Greggor, A.L., Plotnik, J.M., 2021. The challenges of replicating research on endangered species. Anim. Behav. Cogn. 8, 240–246.

Stropp, J., Ladle, R.J., Emilio, T., Lessa, T., Hortal, J., 2022. Taxonomic uncertainty and the challenge of estimating global species richness. J. Biogeogr. 49, 1654–1656.

Studer, A., Thieltges, D.W., Poulin, R., 2010. Parasites and global warming: net effects of temperature on an intertidal host-parasite system. Mar. Ecol. Prog. Ser. 415, 11–22.

Tam, J., Lagisz, M., Cornwell, W., Nakagawa, S., 2022. Quantifying research interests in 7521 mammalian species with *h*-index: a case study. GigaScience 11, giac074.

Titley, M.A., Snaddon, J.L., Turner, E.C., 2017. Scientific research on animal biodiversity is systematically biased towards vertebrates and temperate regions. PLoS One 12, e0189577.

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

Westoby, M., 2002. Choosing species to study. Trends Ecol. Evol. 17, 587.

Wuchty, S., Jones, B.F., Uzzi, B., 2007. The increasing dominance of teams in production of knowledge. Science 316, 1036–1039.