BIOLOGY LETTERS

Research

CrossMark
click for updates

Author for correspondence:
Robert Poulin
e-mail: robert.poulin@otago.ac.nz

THE ROYAL SOCIETY PUBLISHING

Evolutionary biology

Taxonomic distribution of cryptic diversity among metazoans: not so homogeneous after all

Gerardo Pérez-Ponce de León[1] and Robert Poulin[2]

[1]Departamento de Zoología, Instituto de Biología, Universidad Nacional Autónoma de México, Ciudad Universitaria, México D.F., Mexico
[2]Department of Zoology, University of Otago, Dunedin, New Zealand

RP, 0000-0003-1390-1206

Cryptic diversity plagues estimates of biodiversity, conservation efforts and attempts to control diseases and invasive species. Here, we re-visit a decade-old assessment of whether or not cryptic species are homogeneously reported among higher metazoan taxa. We compiled information from an extensive survey of the literature to recover all reports of cryptic species among metazoans. After correcting for currently known species richness and research effort per taxon, we find that cryptic species are over-reported in some taxa and under-reported in others. Although several taxa showing either a lack or an excess of reported cryptic species were poorly studied invertebrate groups, we found that cryptic species were over-reported in amphibians, reptiles and crustaceans, all relatively well-studied groups. The observed heterogeneity in the distribution of reported cryptic species may reflect taxon-specific properties affecting either the propensity for cryptic species to be formed or their likelihood of being detected by conventional taxonomy. Either way, the implications of cryptic diversity may not apply equally across all taxonomic groups.

1. Introduction

The growing recognition that cryptic species, i.e. genetically distinct but morphologically similar species, are ubiquitous requires estimates of biodiversity to be recalibrated [1,2]. Cryptic species have also huge implications for areas of applied ecology where accurate species identification is crucial, such as the conservation of rare species, management of invasive species and control of diseases [3–5]. Yet we have a limited understanding of the distribution of cryptic species, in particular whether they occur more frequently in certain higher taxa than others.

A decade ago, two widely cited articles [4,6] issued competing hypotheses regarding the distribution of cryptic species among taxa. Bickford *et al.* [4] used a literature survey to show that reports of cryptic species were published at an increasing rate, and they argued that cryptic species were heterogeneously distributed. For example, based on their biological properties, they expected frogs and arthropods to contain more cryptic species than other taxa. They provided no rigorous test of this prediction, however. Using a similar dataset on published cryptic species reports (CSR), Pfenninger & Schwenk [6] showed instead that cryptic species were homogeneously reported among taxa. They argued that for a given known species richness and correcting for study effort, reports of cryptic species are evenly distributed among taxa. They identified taxa falling above or below the confidence intervals of their regressions, however, suggesting an excess of CSR for some taxa and a deficit for others. Pfenninger & Schwenk [6] were unfairly criticized for combining non-monophyletic groups and obscuring

the variation among taxa in numbers of CSR through a log-transformation [7]. Nevertheless, given the contrasting conclusions of these two influential papers [4,6], whether or not reported cryptic diversity is homogeneously distributed among taxa remains unclear.

With another decade of data accumulated since those landmark publications, it is time to re-assess the two competing hypotheses. Moreover, we see a potential flaw in the way information was retrieved from the literature in these studies [4,6] that may affect their conclusions. We question the validity of defining a cryptic species report as a publication that uses the term 'cryptic species' (or 'sibling species') in its title, abstract or keywords. Many such publications refer to the phenomenon without actually applying molecular methods to uncover genetic variation among individuals that are morphologically indistinguishable. This may lead to inflated numbers that do not properly reflect the discovery of cryptic diversity in various taxa.

Here, we re-examine the distribution of CSR among metazoan taxa. We first demonstrate that including any publication mentioning 'cryptic (or sibling) species' misrepresents the true number of CSR. We then use the numbers of true CSR per taxon to test whether they are homogeneously distributed among taxa based on currently known species richness and research effort per taxon. Our findings reveal certain higher taxa in which cryptic species are either over- or under-reported, hinting at taxonomic heterogeneity in the discovery of cryptic species.

2. Material and methods

We conducted an exhaustive search of the ISI Web of Science™ for the period 1978–2015, to gather data from studies on cryptic species as in [6], with a strict focus on metazoans (see the electronic supplementary material, S1 for search details). Each record was individually checked to eliminate non-relevant articles, and distinguish between 'true' CSR and papers referring to cryptic species without the authors conducting a search for cryptic species. True CSR are studies that present molecular data (allozyme electrophoresis, random amplified polymorphic DNA, DNA sequencing, etc.) resulting in the recognition of cryptic species, either as their main purpose or as a by-product. We retained studies where a specific search for cryptic species was conducted but none was found. We recorded both true CSR and articles that just mentioned cryptic species for the period 1978–2006, to allow a contrast with Bickford et al.'s [4] analysis covering the same period; from 2007 onwards, we only recorded true CSR. The final dataset includes 2560 articles: 1441 CSR (1978–2015) and 1119 other articles (1978–2006) just mentioning cryptic species (electronic supplementary material, S2). Both annual numbers of CSR and of articles mentioning cryptic species were plotted against year of publication, as in [4], to determine how patterns differed based on how a cryptic species report is defined.

For each CSR, we recorded the year of publication, higher taxon concerned and the number of cryptic species detected beyond the originally known species (electronic supplementary material, S2). Data were summed for each of 34 higher taxa, representing mostly phyla but also classes in the case of diverse groups (table 1). We used three measures of the discovery rate of cryptic diversity: (i) number of CSR, (ii) total number of cryptic species found in those reports and (iii) mean number of cryptic species per CSR. Although involving a genetic search for cryptic species, a few CSR found none; excluding these from the number of CSR per taxon did not affect the results.

For each taxon, we obtained the known extant species richness from the Species 2000 website (http://www.catalogueoflife.org/col/browse/tree), and the total number of articles on that taxon from a search of the ISI Web of Science™ database (e.g. for Annelida, search term = Annelid*; table 1). All variables were log-transformed prior to analysis.

As in [6], we regressed the three above measures of cryptic diversity against species richness across all 34 taxa. Residuals from these regressions represent variation among taxa in the rate at which cryptic diversity is reported, corrected for species richness. Since the mean number of cryptic species per CSR showed no relation with taxon species richness ($p = 0.853$), it is not considered further. We then regressed the total number of articles per taxon against species richness ($F_{1,32} = 51.95$, $R^2 = 0.619$, $p < 0.0001$), with the residuals providing a measure of taxonomic study bias. Positive residuals indicate greater study effort than expected for a given species richness, and negative ones indicate lower than expected study effort. Finally, still as in [6], we regressed the residual measures of cryptic diversity (see above) against taxonomic study bias. Visual inspection suggested possible curvilinearity, which was tested by running second-order polynomial regressions and assessing the significance of the quadratic term.

3. Results

Annual numbers of CSR and of other articles simply mentioning cryptic species were strongly correlated with each other (see the electronic supplementary material, S1). In addition, both measures increased exponentially over time. However, the rise in the number of articles simply mentioning cryptic species was much steeper than that of CSR, and the gap between them got larger over time (electronic supplementary material, S1). This shows that counting any article simply mentioning cryptic species leads to grossly inflated estimates of cryptic species reporting.

Our search produced 1441 'true' CSR, which uncovered 4083 cryptic species. Both the number of CSR ($F_{1,32} = 11.49$, $R^2 = 0.584$, $p < 0.0001$) and the total number of cryptic species found in those reports ($F_{1,32} = 43.32$, $R^2 = 0.575$, $p < 0.0001$) covaried with taxon species richness. Some taxa fell outside the regression's 95% CIs (figure 1). Residuals of both these regressions were significantly related to taxonomic study bias (figure 2). In neither case was the quadratic term significant in a polynomial regression, therefore both residual measures of cryptic diversity were best related to taxonomic study bias by a linear regression (number of CSR: $F_{1,32} = 15.62$, $R^2 = 0.328$, $p = 0.0004$; total number of cryptic species found: $F_{1,32} = 11.59$, $R^2 = 0.266$, $p = 0.0018$). In both relationships, some taxa fell outside the 95% CIs (figure 2). Thus, for a given species richness and study effort, Platyhelminthes, Crustacea, Onychophora, Cycliophora, Reptilia and (to a lesser extent) Amphibia show an excess of CSR and cryptic species found, whereas Collembola, Diplopoda, Acanthocephala and Cephalochordata show a deficit in CSR and cryptic species found.

4. Discussion

Cryptic species confound estimates of biodiversity, conservation efforts and attempts to manage undesirable species [1–5]. We found that overall the number of cryptic species reported is significantly and linearly related to the number of

Table 1. Numbers of CSR, actual cryptic species found, known species richness and study effort based on articles published for each higher taxon.

taxon	CSR	CSR finding no cryptic species	no. cryptic species	total known species richness[a]	total no. articles in Web of Science™ (1978–2015)
Acanthocephala	3	0	8	1192	2138
Amphibia	99	0	312	6439	44 961
Annelida	50	5	128	17 210	5881
Arachnida	60	5	150	110 615	3188
Ascidiacea	10	1	23	2252	571
Aves	36	3	117	10 360	10 668
Bivalvia	26	1	75	8838	22 811
Bryozoa	14	0	42	5486	4369
Cephalochordata	1	0	2	30	631
Cephalopoda	9	0	13	585	5752
Chaetognatha	4	0	9	179	852
Cnidaria	40	5	87	10 105	5092
Collembola	1	0	1	9734	4243
Crustacea	159	8	631	80 122	39 322
Cycliophora	2	0	4	2	74
Diplopoda	3	0	6	7753	1420
Echinodermata	27	1	81	7509	7488
Fish[b]	146	8	468	32 838	328 774
Gastropoda	75	8	171	30 239	19 063
Gastrotricha	1	0	6	790	360
Insecta	300	28	735	1 013 825	202 670
Mammalia	85	6	183	5825	351 382
Merostomata	1	0	1	5	16
Nematoda	61	6	109	24 773	58 949
Nematomorpha	1	0	7	351	224
Nemertea	7	0	20	1200	763
Onychophora	7	0	30	182	419
Platyhelminthes	71	4	187	29 285	3133
Porifera	17	2	42	8346	2589
Pycnogonida	2	0	5	1322	376
Reptilia	99	2	234	9789	18 995
Rotifera	19	1	147	1583	6033
Sipuncula	4	0	11	320	645
Tardigrada	1	0	38	1157	1234

[a]From Species 2000 website, Catalogue of Life (http://www.catalogueoflife.org/col/browse/tree), updated 25 March 2016.
[b]Fish = Actinopterygii + Elasmobranchii + Holocephali + Sarcopterygii.

species per taxon after correcting for study bias. However, we show that after adjusting for variation in species richness and research effort, cryptic species are over-reported in some taxa and under-reported in others. This suggests taxonomic heterogeneity in the occurrence of cryptic species among metazoans.

We initially used three measures of the discovery rate of cryptic diversity: the number of CSR, the total number of cryptic species found and the mean number of cryptic species per CSR. The latter measure varied among taxa [7], but did not correlate with taxon species richness (nor with taxonomic study bias: $p > 0.10$). It is unclear which of the other two measures best captures the discovery rate of cryptic species. However,

with some exceptions, they both identify many of the same taxa lying either above or below the confidence interval in the regression against taxonomic study bias. Interestingly, these taxa are not the same as those found in [6], possibly owing to the extra decade of data incorporated in our analysis, or because we used a stricter definition of CSR. Regardless, Pfenninger & Schwenk [6] found that cryptic species were under-reported in most mollusc groups, and over-reported in mammals and amphibians. Apart from a similar result for amphibians, we instead found that cryptic species in those other groups were reported roughly as frequently as expected based on species richness and study effort.
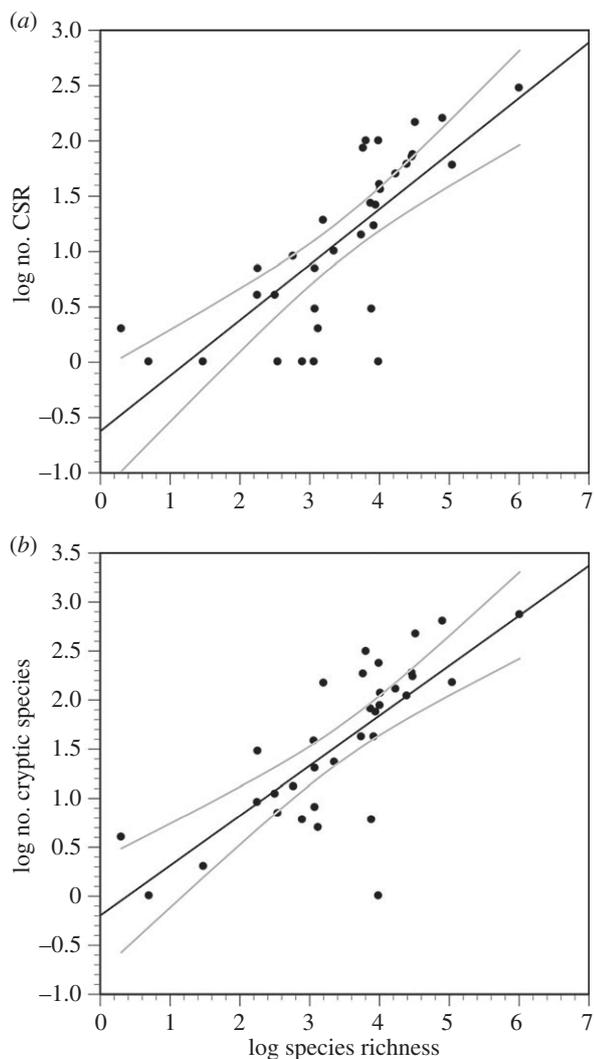
Figure 1. Number of CSR per taxon (*a*) and total number of cryptic species found (*b*) against known extant species richness, across 34 higher taxa. Grey lines encompass the 95% CIs.

Figure 2. Residual number of CSR per taxon (*a*) and residual total number of cryptic species found (*b*) against taxonomic study bias, across 34 higher taxa. Residuals are from the regressions in figure 1. Grey lines encompass the 95% CIs. In addition to the labelled points, those for insects (I), fish (F), birds (B) and mammals (M) are identified.

Many taxa falling outside the 95% confidence intervals were invertebrate groups that have received little attention; therefore, their apparent excess or lack of CSR should be taken with caution. However, this does not apply to amphibians, reptiles and crustaceans: all showed an excess of reported cryptic species after correcting for species richness and research effort. Why are cryptic species particularly common in those taxa? Ecological differences between habitats can affect rates of diversification. For instance, the discontinuity of freshwater habitats and the greater primary productivity of terrestrial environments may explain the disproportionately higher diversification of freshwater and terrestrial organisms relative to marine ones [8,9]. This cannot explain the patterns we found, as many terrestrial or freshwater taxa did not deviate from the expected number of CSR. Intrinsic properties of organisms may also influence the occurrence of cryptic species. For example, taxa like amphibians that rely on non-morphological characters for species recognition may harbour much cryptic diversity missed by earlier taxonomic studies [4]. In addition, cryptic species should be rare among large-bodied taxa, because their dispersal abilities prevent local genetic differentiation [10]. However, amphibians, reptiles and crustaceans are not the smallest taxa considered here.
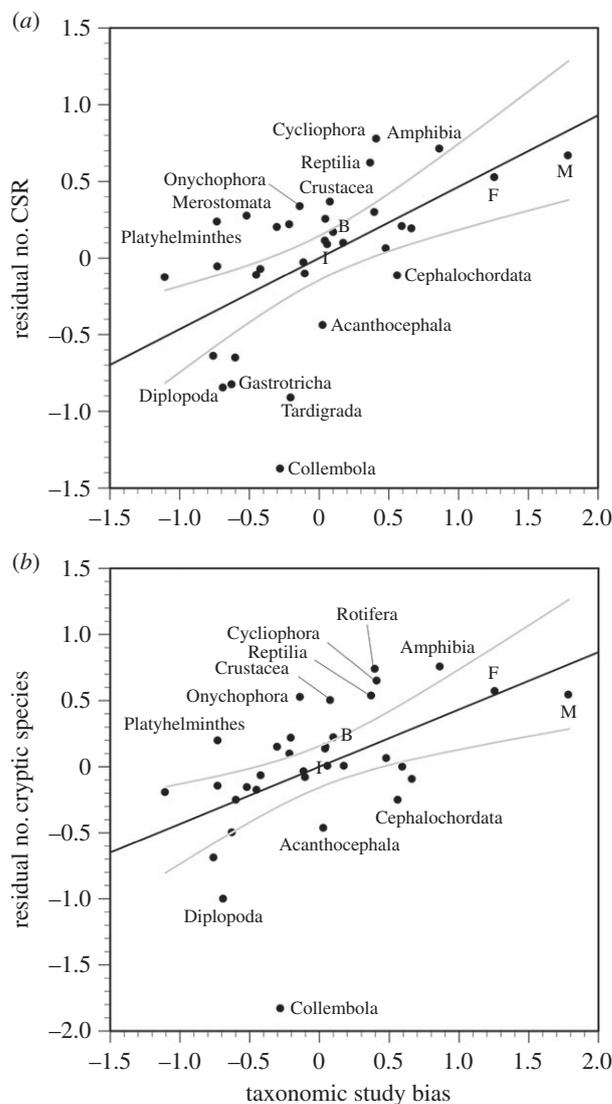
Although the underlying reasons remain unclear, there is heterogeneity among taxa in the reporting of cryptic diversity after accounting for variation in species richness and research effort. Whether cryptic species represent the limits of the discriminatory power of traditional taxonomic approaches, or a genuine facet of biodiversity arising from recent speciation events, identifying higher taxa in which they occur more frequently can focus future biodiscovery research.

References

1. Scheffers BR, Joppa LN, Pimm SL, Laurance WF. 2012 What we know and don't know about Earth's missing biodiversity. *Trends Ecol. Evol.* **27**, 501–510. (doi:10.1016/j.tree.2012.05.008)

2. Loxdale HD, Davis BJ, Davis RA. 2016 Known knowns and unknowns in biology. *Biol. J. Linn. Soc.* **117**, 386–398. (doi:10.1111/bij.12646)

3. Witt JDS, Threloff DL, Hebert PDN. 2006 DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation. *Mol. Ecol.* **15**, 3073–3082. (doi:10.1111/j.1365-294X.2006.02999.x)

4. Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I. 2007 Cryptic species as a window on diversity and conservation. *Trends Ecol. Evol.* **22**, 148–155. (doi:10.1016/j.tree.2006.11.004)

5. Nadler SA, Pérez-Ponce de León G. 2011 Integrating molecular and morphological approaches for characterizing parasite cryptic species: implications for parasitology. *Parasitology* **138**, 1688–1709. (doi:10.1017/S003118201000168X)

6. Pfenninger M, Schwenk K. 2007 Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evol. Biol.* **7**, 121. (doi:10.1186/1471-2148-7-121)

7. Trontelj P, Fiser C. 2009 Cryptic species diversity should not be trivialised. *Syst. Biodiv.* **7**, 1–3. (doi:10.1017/S1477200008002909)

8. Grosberg RK, Vermeij GJ, Wainwright PC. 2012 Biodiversity in water and on land. *Curr. Biol.* **22**, pR900-R903. (doi:10.1016/j.cub.2012.09.050)

9. Wiens JJ. 2015 Faster diversification on land than sea helps explain global biodiversity patterns among habitats and animal phyla. *Ecol. Lett.* **18**, 1234–1241. (doi:10.1111/ele.12503)

10. Beheregaray LB, Caccone A. 2007 Cryptic biodiversity in a changing world. *J. Biol.* **6**, 9. (doi:10.1186/jbiol60)