Received: 12 December 2021 | Revised: 22 August 2022 | Accepted: 25 August 2022

DOI: 10.1111/geb.13589

META-ANALYSIS

Global Ecology and Biogeography A Journal of Macroecology

WILEY

Avoiding growing pains in reproductive trait databases: the curse of dimensionality

Samuel C. Ginther | **Hayley Cameron** | **Craig R. White** | **Dustin J. Marshall**

Centre for Geometric Biology, Monash University, Clayton, Victoria, Australia

Correspondence
Samuel C. Ginther, Centre for Geometric Biology, Monash University, 18 Innovation Walk, Clayton, VIC 3800, Australia.
Email: samuel.ginther@monash.edu

Handling Editor: Franziska Schrodt

[Correction added on 16 September 2022, after first online publication: article title 'Avoidance of growing pains in reproductive trait databases: The curse of dimensionality' has been corrected to read 'Avoiding growing pains in reproductive trait databases: the curse of dimensionality'.]

Abstract

Aim: Reproductive output features prominently in many trait databases, but the metrics describing it vary and are often untethered to temporal and volumetric dimensions (e.g., fecundity per bout). The use of such ambiguous reproductive measures to make broad-scale comparisons across taxonomic groups will be meaningful only if they show a 1:1 relationship with a reproductive measure that explicitly includes both a volumetric and a temporal component (i.e., reproductive mass per year). We sought to map the prevalence of ambiguous and explicit reproductive measures across taxa and to explore their relationships with one another to determine the cross-compatibility and utility of reproductive metrics in trait databases.

Location: Global.

Time period: 1990–2021.

Major taxa studied: We searched for reproductive measures across all Metazoa and identified 19,785 vertebrate species (Chordata), and 440 invertebrate species (Arthropoda, Cnidaria or Mollusca).

Methods: We included 37 databases, from which we summarized the commonality of reproductive metrics across taxonomic groups. We also quantified scaling relationships between ambiguous reproductive traits (fecundity per bout, fecundity per year and reproductive mass per bout) and an explicit measure (reproductive mass per year) to assess their cross-compatibility.

Results: Most species were missing at least one temporal or volumetric dimension of reproductive output, such that reproductive mass per year could be reconstructed for only 4,786 vertebrate species. Ambiguous reproductive measures were poor predictors of reproductive mass per year; in no instance did these measures scale at 1:1.

Main conclusions: Ambiguous measures systematically misestimate reproductive mass per year. Until more data are collected, we suggest that researchers should use the clade-specific scaling relationships provided here to convert ambiguous reproductive measures to reproductive mass per year.

KEYWORDS
database, reproduction, reproductive allometry, reproductive frequency, reproductive trait correlations, systematic map, trait incompatibility

1 | INTRODUCTION

Online databases of biological traits have proliferated - in the Open Traits Network (https://opentraits.org/), 35 of the reported 49 databases have been published since 2015. The compilation of such data has been a tremendous effort and has helped synthesize information for a wider breadth of researchers, while also unlocking new, broad-scale questions. For example, PanTHERIA (Jones et al., 2009), a database composed of >5,000 mammal species, has played a key role in the identification of new patterns of biodiversity (Newbold et al., 2015), extinction risk (Fritz et al., 2009) and animal movement (Tucker et al., 2018). Likewise, a 30-year-old fish database, FishBase (Froese & Pauly, 2021), has provided access to a host of fundamental and applied questions that were previously inaccessible. We are eager advocates for, and frequent users of, biological databases (Barneche et al., 2018; White et al., 2019, 2021), but as these databases grow, so do the potential challenges and risks of using them (Mills et al., 2015; Salguero-Gómez et al., 2021; Wilkinson et al., 2016).

Reproduction features prominently in many trait databases and is important for understanding short-term population dynamics, persistence, fitness and invasiveness (Jelbert et al., 2019; Kokko, 2021). Yet the metrics used to describe reproduction are variable and show taxonomic idiosyncrasies; even those working in the same field and on similar taxa use different metrics. For example, reproduction in birds is often described using clutch size (i.e., fecundity per bout; Myhrvold et al., 2015), whereas mammal reproduction is often described using total litter mass (i.e., reproductive mass per bout; Blueweiss et al., 1978). Nonetheless, many analyses seek to understand how reproduction varies across clades and life histories, and for such comparisons to be made, a common currency is essential. For example, studies that mix total brood mass and offspring number will not be comparable if individual offspring mass also shows significant variation among species, which it does (reviewed by Marshall et al., 2018).

Although the above examples demonstrate inconsistencies in volumetric measurements of reproduction, variation in temporal dimensions of reproduction also exists. Some species produce only one clutch per reproductive season, whereas others reproduce almost every day (Healy et al., 2019; Marshall et al., 2021). Many studies report reproduction as fecundity per bout, but far fewer report the number of bouts per unit time (Healy et al., 2019; Marshall et al., 2021; Meiri et al., 2012). Overall, the scope for making meaningful comparisons of reproduction across clades depends on consistency in the dimensionality of metrics reported. Yet the prevalence of these issues across animal trait databases remains unclear, as do their consequences for broad-scale comparisons.

For simplicity, we can classify reproductive measures into two broad groups: "ambiguous reproductive measures" are those untethered from temporal and volumetric dimensions (e.g. fecundity per bout); or those that include one dimension, but not the other (for a full list, see Table 1). In contrast, "explicit reproductive measures" must be anchored in both mass and time - thus, reproductive mass per year is the only reproductive trait that is truly explicit (Table 1). Reproductive mass per year requires much more information: offspring have to be weighed and counted, and reproduction needs to be assessed for an entire reproductive season or lifetime (Figure 1a). Consequently, we suspect that few compilations will include reproductive mass per year, although this metric is both the most biologically meaningful, because it captures a rate of resource investment, and comparable, because it allows comparisons among divergent clades with very different life histories (e.g., Ernest et al., 2003; Hatton et al., 2019).

Ambiguous reproductive measures can still be used to make direct comparisons of reproduction across species if their relationship with reproductive mass per year is 1:1 - in other words, positively isometric. However, ambiguous reproductive measures will provide reliable estimates of the explicit reproductive measure only in a very restrictive set of conditions: if the underlying ambiguous components that make up reproductive

TABLE 1 Reproductive traits identified in our systematic map, with their descriptions and whether they are ambiguous or explicit measures of reproductive output

Trait	Description	Dimensions
Offspring size	Size of offspring [mass (in grams) or length (in millimetres)]	Ambiguous
Reproductive frequency	Number of times an adult reproduces throughout a year (i.e., number of reproductive events per year)	Ambiguous
Fecundity per bout	Total number of all offspring produced during a single reproductive event	Ambiguous
Fecundity per year	Total number of all offspring produced during a year	Ambiguous
Reproductive mass per bout	Total mass of all offspring produced during a single reproductive event (in grams)	Ambiguous
Reproductive mass per year	Total mass of all offspring produced during a year (in grams per year)	Explicit

Note: The fecundity per year was calculated from the fecundity per day for 81 arthropods (exclusively Hymenopterans) (Jervis & Ferns, 2011).

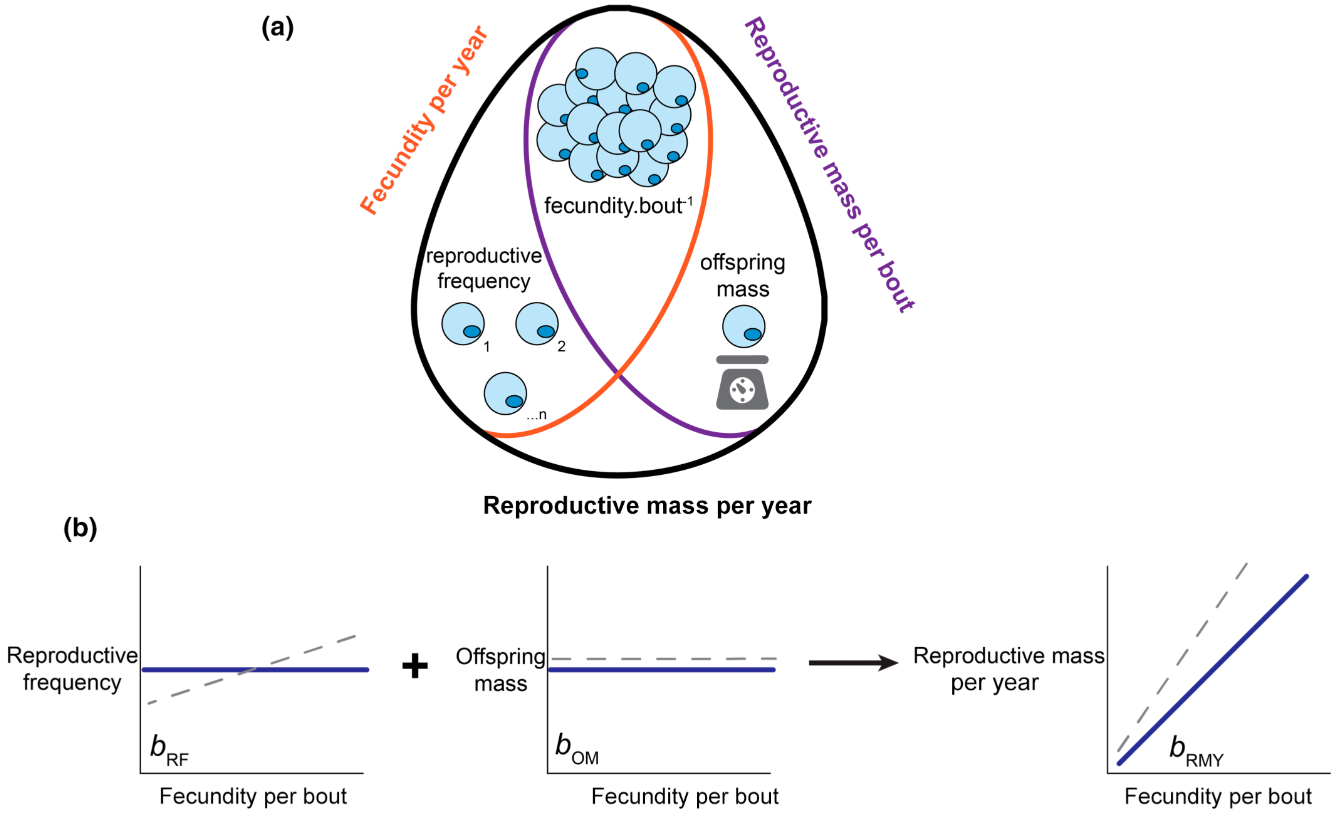


FIGURE 1 (a) Reproductive mass per year (in grams per year) is a combination of fecundity per bout (as number of offspring per event), offspring mass (in grams) and reproductive frequency (as reproductive events per year) (black outline). Reproductive traits used to calculate fecundity per year (as number of offspring per year) and reproductive mass per bout (in grams per bout) indirectly are shown in orange and purple, respectively. (b) Scaling relationships between the ambiguous reproductive traits that comprise reproductive mass per year: Fecundity per bout and reproductive frequency (b_{RF}; left panel); and fecundity per bout and offspring mass (b_{OM}; middle panel). The values of b_{RF} and b_{OM} will yield the scaling relationship between fecundity per bout and reproductive mass per year (b_{RMY}) via the equation $b_{RMY} = b_{RF} + b_{OM} + 1$ (right panel). Accordingly, $b_{RMY} = 1$ under the restrictive circumstances of both b_{RF} and $b_{OM} = 0$ (blue line) or b_{RF} and b_{OM} are the perfect inverse of each other (e.g., $b_{RF} = -0.75$ and $b_{OM} = +0.75$; not shown). Any other combinations of b_{RF} or b_{OM} will result in $b_{RMY} \neq 1$, as demonstrated in panel (b), where $b_{RMY} > 1$ when $b_{RF} > 0$ and $b_{OM} = 0$ (dashed grey line). Note that other values of b_{RMY} are plausible and depend on b_{RF} and b_{OM}

mass per year do not covary with one another (Figure 1, blue line). For example, if fecundity per bout shows no covariance with offspring mass or reproductive frequency, positive isometry will occur between fecundity per bout and reproductive mass per year (Figure 1b, blue line). But such covariances between reproductive traits are well documented among species: species that produce many offspring per bout often produce relatively smaller offspring (Burger et al., 2019; Lack, 1947; Thorson, 1950), whereas species with lower reproductive outputs per bout tend to reproduce more often in any one season (Marshall et al., 2021). Thus, comparisons of reproductive output that lack volumetric and/or temporal dimensions are likely to be an inaccurate depiction of reproductive effort. We note, however, that in instances where relationships deviate from isometry, ambiguous traits might provide appropriate proxies for reproductive mass per year if, and only if, the nonlinear scaling relationships between these variables are known (Figure 1b, grey line). Nonetheless, explorations of such scaling relationships are entirely lacking. A necessary step is to assess formally whether temporally and volumetrically ambiguous reproductive measures are synonymous to an explicit one - that is, reproductive mass per year - so that reproductive effort can be compared properly across broad taxonomic clades.

Here, we used a systematic map, a transparent process to collate and describe existing data (James et al., 2016), to identify animal databases or compilations that report reproductive traits. Our primary goals were: (1) to determine the commonality of reproductive metrics used across major animal groups; and (2) to asses how well ambiguous reproductive measures approximated reproductive mass per year. Our intention was to inform the future collection and compilation of reproductive traits to enable researchers to make broader taxonomic comparisons of reproductive output.

2 | METHODS

We followed the guidelines of the systematic mapping methodology (James et al., 2016) to determine what reproductive traits are provided in animal databases and to quantify their relationships with one another.

2.1 | Review team

One primary reviewer and three stakeholders set the scope and question of the systematic map; set the inclusion and exclusion criteria; aided in including or excluding databases; and helped to describe the findings (James et al., 2016).

2.2 | Scoping and searching

We searched the literature from 2020 to 2021; we conducted the final search in October 2021. Initially, to identify literature to screen, we trialled a combination of search terms (Supporting Information Table S1) in the publication database, Web of Science. After trialing 16 search terms, we selected the search term "(*phylum*) AND (life history* OR trait) AND (database* OR compil*) NOT (plant*)", where *phylum* was substituted each search with each metazoan phylum (for full list of search terms, see Supporting Information Table S1). These search terms resulted in a total of 28,028 hits.

2.3 | Screening

We established inclusion and exclusion criteria before the searching phase to limit any biases introduced by the primary reviewer (Supporting Information Table S2). The main inclusion criteria were that databases must have listed continuous, numerical values of adult body size (mass or length) and either fecundity (total number of offspring per bout or total number of offspring per time) or reproductive mass (total offspring mass per bout or total offspring mass per time); see Table 1 for a description of all reproductive traits identified by our systematic map. Databases that did not include adult size were excluded, given that body size is a major driver of reproductive output and should be accounted for when making comparisons across taxa (Barneche et al., 2018; Ernest et al., 2003; Hatton et al., 2019; Healy et al., 2019; Nørgaard et al., 2021). Additionally, we collected offspring mass to calculate reproductive mass when direct measurements were unavailable. Offspring length was also collected to assess how well it approximated offspring mass (described below). Although we recognize offspring size as an important component of reproductive investment, we did not include databases that included only offspring size, because such data provide no indication of reproductive output in the absence of fecundity; only one study was excluded for not meeting this criterion (Neuheimer et al., 2015). We excluded categorical traits that were reported as a range of values (e.g., fecundity: 50–500), and only three databases reported traits in this way (Dziock et al., 2011; Faulwetter et al., 2014; Salguero-Gómez et al., 2016). We also excluded data that were missing taxonomic information, because comparative analyses would not be able to account properly for phylogeny; only one database did not meet this criterion (Sarremejane et al., 2020).

To identify the eligible databases returned by our literature searches (as previously described), we sorted the titles by "relevance" and screened titles and abstracts of the first 500 hits of each phylum-specific search term. If studies appeared to match the inclusion criteria, they were marked and uploaded into the Rayyan Systematic Reviews web application (Ouzzani et al., 2016), where each study and its associated database was reviewed fully. In total, 3,410 titles and abstracts were reviewed in Web of Science, and 240 studies were reviewed fully. Studies were sorted into two groups, "include" or "exclude", fully assessing them after; a list of assessed articles can be found in the data archived with Dryad and online supplemental data. Screening and eligibility assessments reduced the number of eligible databases to 42 (Supporting Information Figure S1).

2.4 | Coding

We coded 42 databases into MICROSOFT EXCEL (v.16.46), noting: (1) reference information, (2) species information, and (3) trait data. For each species in each online database, we coded the species name and the following reproductive traits (when available): adult body size (mass or length); fecundity measure (fecundity per bout, fecundity per time, reproductive mass per bout or reproductive mass per time); offspring size (mass or length); and reproductive frequency (number of reproductive events per time) (Table 1). We also extracted the numerical values for each trait to explore relationships between ambiguous and explicit reproductive measures. When not reported directly, we used different reproductive trait combinations to calculate unreported reproductive traits. For example, fecundity as a rate (i.e., fecundity per time) was calculated by multiplying fecundity per bout and reproductive frequency (Figure 1a). Note that "per time" can refer to "per year" or "per day", but can be converted easily into a common currency. We converted all measures of fecundity per time and reproductive mass per time to a yearly rate, and we refer to these hereafter as fecundity per year or reproductive mass per year, respectively, unless specified otherwise (Table 1). Additionally, only offspring mass (and not offspring length) was used to calculate reproductive mass (per bout and per year), because we wished to minimize error that can occur from using length-to-mass conversions that are not species specific (Feldman & Meiri, 2013; Froese & Pauly, 2021; Meiri, 2010; Neuheimer et al., 2015).

After recording the observations contained within all eligible databases into a single EXCEL file, we used RSTUDIO (v.1.4.1106; R Core Team, 2018) and "tidyverse" packages (v.1.3.1) (Wickham et al., 2019) to summarize and combine duplicate species observations. In some instances, multiple sources are listed for each species because one trait might have been identified in one database, but not another; that is, traits reported with multiple dimensions (e.g., fecundity per bout and fecundity per year) for the same species could have originated from different sources. When we found that the same species had duplicate trait observations across multiple databases (e.g., fecundity per bout was reported twice for the same species), we defaulted to the oldest database and removed duplicate observations from the more recent database(s). We omitted duplicate observations to avoid biases in our assessment of the commonality of reproductive measures, given that species from well-represented taxa were found across multiple databases (Supporting Information Table S3). However, if users are interested in combining multiple trait estimates (e.g., to obtain mean trait values for these species), they can refer to the original databases in the Appendix 1 and to our Supporting Information (Table S3). After duplicate species were removed, our final database included observations of reproductive traits for 20,225 species from 37 studies (Appendix 1; Supporting Information Figure S1).

We used the Global Biodiversity Information Facility (GBIF) database within the "taxize" R package (v.0.9.99; Chamberlain & Szöcs, 2013) to resolve species names and identify the upstream phylogenetic nomenclature of all species included in our systematic map. We then checked the taxonomic information for obvious errors (e.g., phyla that were not Metazoa), and we used the "sample" function in base R to generate a random sample of 100 species to verify the accuracy of the taxonomic nomenclature (see data archived with Dryad). One hundred per cent of randomly sampled species had accurate "phylum" and "class" nomenclature, hence no further checks of taxonomy were made.

2.5 | Determining the comparability of reproductive output across taxa

Initially, we summarized the commonality of reproductive measures reported in databases by constructing structural matrices depicting the frequency and distribution of species across animal classes and reproductive traits. Note that proportions calculated are relative to the total number of species of that clade extracted from the databases, and not the global biodiversity estimated for these clades.

2.6 | How well does offspring length approximate offspring mass?

We used linear regression to assess whether offspring length and mass relationships were consistent across animal clades using the "car" package in RStudio (Fox & Weisberg, 2019). To generate linear relationships between offspring mass and length for each animal class, we fitted a linear regression model that included $\log_{10}$-transformed offspring length, animal class (categorical fixed effect), and their interaction ($\log_{10}$-length × class) in a model. The single model generated class-specific intercepts and slope coefficients that had the same structure: $\log_{10}(\text{offspring mass}) = b \times \log_{10}(\text{offspring length}) + \log_{10}(a)$, where b is the slope of the relationship between offspring mass and length, and $\log_{10}(a)$ is the y-intercept. Note that we had data to assess the relationship between offspring length and mass for only three vertebrate classes: Aves, Mammalia and Reptilia.

2.7 | Exploring relationships between reproductive traits across phylogeny

We used linear regression to assess whether ambiguous reproductive traits were good approximates of reproductive mass per year. All class-specific models had the same structure: $\log_{10}(Y) = b \times \log_{10}(X) + \log_{10}(a)$, where X represents ambiguous reproductive traits (fecundity per bout, fecundity per year and reproductive mass per bout), Y represents reproductive mass per year, b is the slope of the relationship between reproductive traits, and $\log_{10}(a)$ is the y-intercept. Ambiguous reproductive traits are synonymous with reproductive mass per year only when their relationship is positively isometric [Figure 1b, blue line; i.e., the slope (b) of the relationship is one, and the y-intercept coefficient (a) is zero]. When the relationship deviates from perfect isometry (i.e., $b \neq 1$ *or* $a \neq 0$), ambiguous measures will misestimate reproductive mass per year (Figure 1b, grey line). We used Wald tests to determine whether class-specific coefficients for slopes (b) and y-intercepts (a) deviated from one and zero, respectively ($p < .05$).

Reproductive measures used to fit these regression models were obtained directly from the original databases or calculated from other reproductive traits; for example, offspring mass multiplied by fecundity per bout yields reproductive mass per bout (Figure 1a). Only offspring mass was used to estimate reproductive traits; we did not use offspring length because offspring length-to-mass conversions can result in additional error in calculations of reproductive output. We had data to assess the relationship between ambiguous reproductive traits and reproductive mass per year for only three vertebrate classes: Aves, Mammalia and Reptilia.

2.8 | Drivers of trait relationships

A condition of positive isometry between ambiguous reproductive measures and reproductive mass per year is that there is no covariance between the underlying ambiguous traits that comprise reproductive mass per year (Figure 1b, blue line). As such, we explored linear-regression models between $\log_{10}$-transformed fecundity per bout and (1) reproductive frequency and (2) offspring mass. To characterize the covariance between fecundity per bout and other reproductive traits formally, we used Wald tests to determine whether slopes and y-intercepts differed from zero for each class ($p < .05$).

3 | RESULTS

Thirty-seven databases included a measure of fecundity (fecundity per bout or fecundity per year) or reproductive mass (reproductive mass per bout or reproductive mass per year) for 20,225 animal species from four phyla (Appendix 1; Supporting Information Table S3). Nearly all species (*c.* 98%) in our dataset were vertebrates (19,785 spp.); the few others were arthropods (199 spp.), cnidarians (11 spp.) and molluscs (230 spp.).

3.1 | Prevalence of reproductive measures across clades

Offspring size was usually reported as mass for the most abundant classes of vertebrates (Figure 2a), but offspring mass was rarely reported for invertebrates. The covariance between offspring mass and length for vertebrates with both measures depended on class, and the scaling of length with mass was always hyperallometric; scaling exponents ranged from 2.3 to 2.9 [$\log_{10}$ offspring length (mm) × vertebrate class: $F_{2,1994} = 13.9$, $p < .001$; Figure 2b; Supporting Information Table S4].

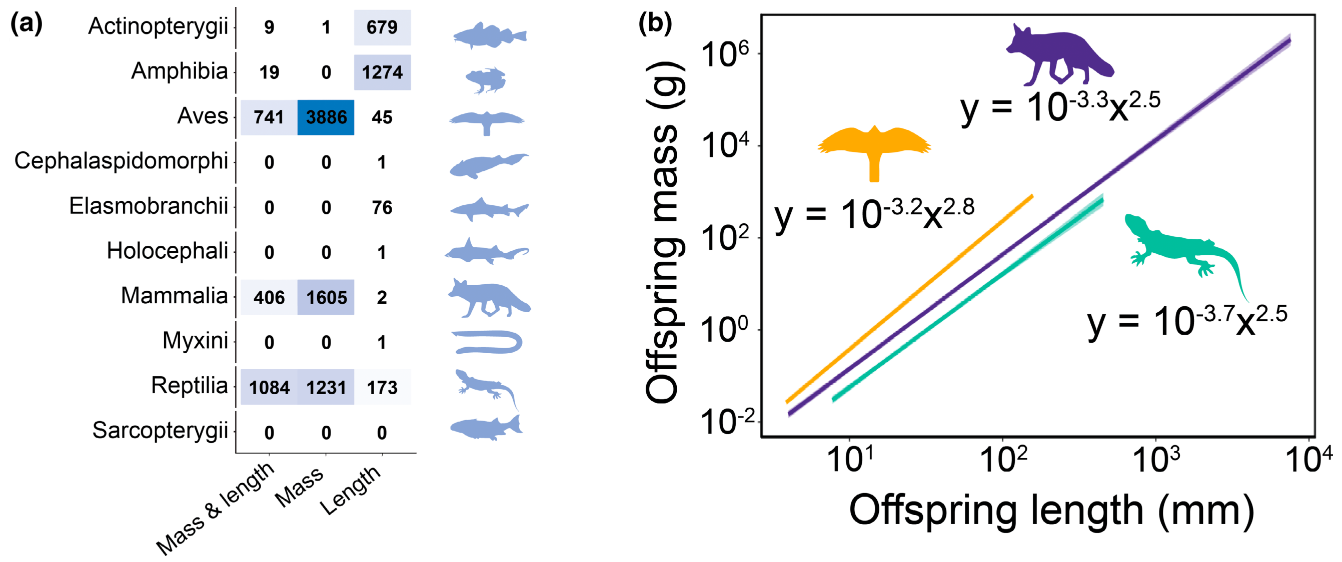
Nearly half of all animals identified by our systematic map had volumetric dimensions of reproductive output. Reproductive mass per bout was rarely reported directly (i.e., 134 Reptilia species), but could be calculated by multiplying offspring mass by fecundity per bout for *c.* 45% of vertebrate species across five clades (8,861 spp.) (Figure 3). For invertebrates, reproductive mass per bout was available for only *c.* 1% of species (Figure 3).

Most species identified by our systematic map were missing temporal dimensions of reproductive output (Figure 3). Fecundity per year was absent from cnidarian and mollusc compilations and was reported directly for *c.* 43% of arthropods (81 insect spp.). Reproductive rate data were relatively more common for vertebrates but were still rare absolutely. Fecundity per year was never reported directly but could be extrapolated for *c.* 36% of vertebrate species across Actinopterygii (one sp.), Amphibia (1,548 spp.), Aves (1,832 spp.), Mammalia (2,267 spp.) and Reptilia (1,448 spp.) by multiplying values of fecundity per bout and reproductive frequency (Figure 3).

Reproductive mass per year, the only metric that allows for meaningful comparisons of reproductive output across species, was rarely reported directly [*c.* 8% of vertebrates (1,527 spp.)] but could be reconstructed from values of fecundity per bout, offspring mass and reproductive frequency for *c.* 16% of vertebrate species (3,259 spp.) (Figure 3). Reproductive mass per year was never reported directly for any invertebrates and could not be reconstructed by combining other reproductive traits.

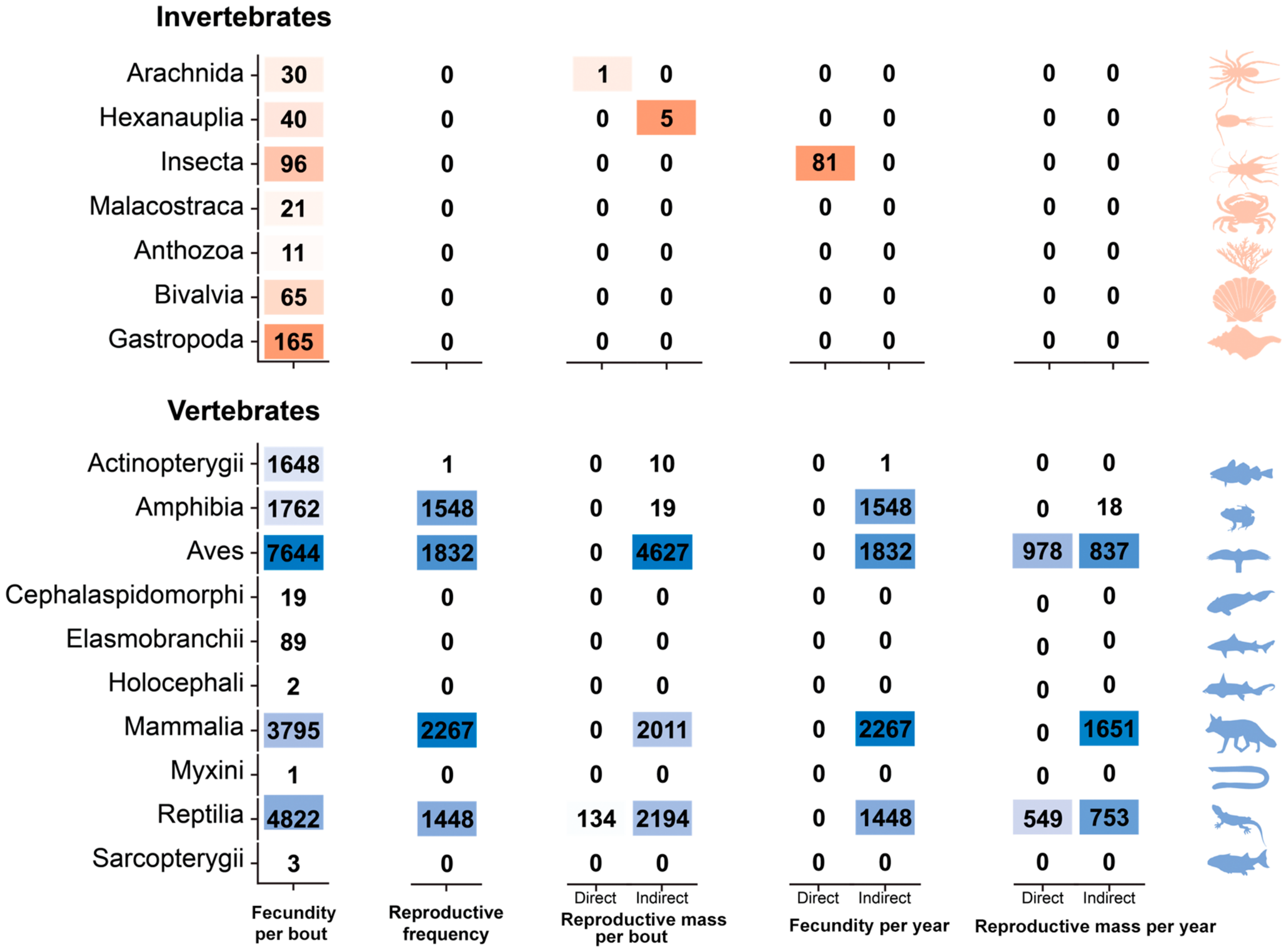
3.2 | Are ambiguous reproductive measures good proxies for reproductive effort?

Ambiguous reproductive measures were never perfect predictors of reproductive mass per year across taxa. Two-thirds of the relationships between ambiguous reproductive measures and reproductive mass per year were positive, but they were always nonlinear (i.e., $b \neq 1$); the remaining trait relationships were negative (Figure 4; Table 2). Nonlinearities between reproductive traits generally



	Mass & length	Mass	Length
Actinopterygii	9	1	679
Amphibia	19	0	1274
Aves	741	3886	45
Cephalaspidomorphi	0	0	1
Elasmobranchii	0	0	76
Holocephali	0	0	1
Mammalia	406	1605	2
Myxini	0	0	1
Reptilia	1084	1231	173
Sarcopterygii	0	0	0

FIGURE 2 (a) Structural matrix of the total number of vertebrate species for which offspring size is reported as "mass & length" simultaneously (left panel), "mass" only (middle panel) and "length" only (right panel). The colour gradient and numbers within cells represent the number of species. (b) The relationship between offspring length and mass for vertebrate species for which both values are reported simultaneously (shadings are ±95% confidence intervals). Equations provide exponents and intercepts for each clade-specific relationship composed of ≥100 species (on the antilogarithmic scale). Silhouettes and colours show major vertebrate clades: Aves (orange), Mammalia (purple) and Reptilia (green)



Invertebrates

	Fecundity per bout	Reproductive frequency	Reproductive mass per bout		Fecundity per year		Reproductive mass per year	
			Direct	Indirect	Direct	Indirect	Direct	Indirect
Arachnida	30	0	1	0	0	0	0	0
Hexanauplia	40	0	0	5	0	0	0	0
Insecta	96	0	0	0	81	0	0	0
Malacostraca	21	0	0	0	0	0	0	0
Anthozoa	11	0	0	0	0	0	0	0
Bivalvia	65	0	0	0	0	0	0	0
Gastropoda	165	0	0	0	0	0	0	0

Vertebrates

	Fecundity per bout	Reproductive frequency	Reproductive mass per bout		Fecundity per year		Reproductive mass per year	
			Direct	Indirect	Direct	Indirect	Direct	Indirect
Actinopterygii	1648	1	0	10	0	1	0	0
Amphibia	1762	1548	0	19	0	1548	0	18
Aves	7644	1832	0	4627	0	1832	978	837
Cephalaspidomorphi	19	0	0	0	0	0	0	0
Elasmobranchii	89	0	0	0	0	0	0	0
Holocephali	2	0	0	0	0	0	0	0
Mammalia	3795	2267	0	2011	0	2267	0	1651
Myxini	1	0	0	0	0	0	0	0
Reptilia	4822	1448	134	2194	0	1448	549	753
Sarcopterygii	3	0	0	0	0	0	0	0

FIGURE 3 Structural matrix of the total number of invertebrate (orange; top panel) and vertebrate (blue; bottom panel) species for which reproductive traits were reported (for trait descriptions, see Table 1). Reproductive traits were measured directly ("direct") or calculated indirectly by combining values for other reproductive traits ("indirect") (see Figure 1a). The colour gradient and numbers within cells represent the number of species for each trait and class combination. Note that there is no overlap of species between "direct" and "indirect" columns

resulted in ambiguous reproductive measures underestimating the interspecific variation in reproductive mass per year by more than two orders of magnitude (Supporting Information Table S5).

Reproductive mass per bout was the least misleading proxy for reproductive mass per year. The scaling of reproductive mass per year with reproductive mass per bout was the closest to isometry and was also less variable across clades: Aves and Mammalia showed the same scaling exponent for this relationship ($b = 0.89$), whereas scaling across Reptilia was slightly steeper ($b = 0.95$) ($\log_{10}$reproductive mass per bout × vertebrate class: $F_{2,4537} = 30.0$, $p < .001$; Figure 4c; Table 2; Supporting Information Table S6).

3.3 | Temporal and volumetric drivers of reproductive trait relationships

Reproductive frequency and offspring mass both covary with fecundity per bout, such that fecundity per bout systematically misestimates reproductive mass per year across vertebrates (Figure 5; Supporting Information Table S7). Aves and Mammalia that produce many offspring per bout also reproduce more often and produce larger offspring than those that produce few offspring per bout. The opposite occurs across Reptilia: species that produce many offspring per bout reproduce less often and produce smaller offspring in comparison to reptiles that produce fewer offspring per bout. Overall then, the covariance between underlying reproductive traits that compose reproductive mass per year (i.e., fecundity per bout, reproductive frequency and offspring mass) result in ambiguous reproductive measures being poor estimates of reproductive mass per year (Figure 4; Table 2).

4 | DISCUSSION

Our capacity to compare species, within and among databases, is only as good as the commonality, or at least compatibility, of the

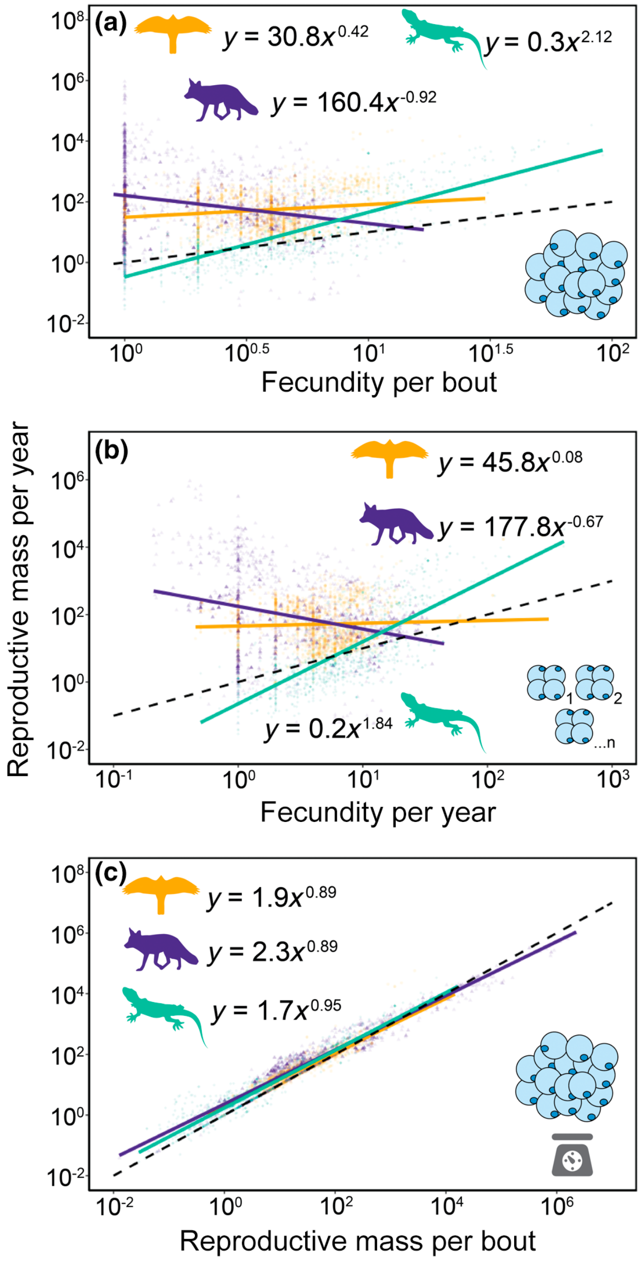


FIGURE 4 The relationship between reproductive mass per year (in grams per year) and: (a) fecundity per bout (as number of offspring per event); (b) fecundity per year (as number of offspring per year); and (c) reproductive mass per bout (in grams per bout). Equations provide exponents and intercepts for each clade-specific relationship composed of ≥100 species (on the antilogarithmic scale). The dashed line shows the null hypothesis for perfect isometry between reproductive traits (i.e., $b = 1$, $a = 0$). Silhouettes and colours are the same as in Figure 2; species of Aves, Mammalia and Reptilia are represented by circles, triangles and diamonds, respectively

trait data therein. We show that ambiguous reproductive measures are the most common metrics of reproduction and that they are poor predictors of reproductive mass per year. For example, fecundity per bout is the most common metric reported and used for vertebrates, yet the relationship between fecundity per bout and reproductive mass per year is negative in mammals; species with the highest fecundities have the lowest rates of total reproductive output. Therefore, our results highlight that any inferences regarding the drivers of reproductive output among species that are based on compilations of ambiguous reproductive measures are likely to be misplaced.

Biologists have long been fascinated by interspecific variation in reproductive output (Barneche et al., 2018; Blueweiss et al., 1978; Hatton et al., 2019; Healy et al., 2019; Savage et al., 2004), but it appears that in most instances we have been underestimating the true level of variation that exists when ambiguous reproductive traits are combined to calculate reproductive mass per year. Our finding of allometric scaling between various ambiguous reproductive measures and reproductive mass per year means that previous attempts to compare reproduction across species have systematically misestimated variation in reproduction. In particular, positive covariance between fecundity per bout and both reproductive frequency and offspring mass increases the variation in reproduction that exists and, ultimately, fails to describe reproductive mass per year. For nearly every ambiguous reproductive measure across vertebrates, interspecific variation in reproductive mass per year was underestimated by other metrics, sometimes by more than three orders of magnitude. If we are to understand variation among different life histories, we must first quantify it accurately; simple measures, such as fecundity per bout, fail to do this.

We believe the best way to rectify inferences about reproduction that are based on ambiguous reproductive measures is to gather the most explicit trait (i.e., reproductive mass per year). In the meantime, however, future studies can use the nonlinear relationships we provide to convert ambiguous measures to approximations of an explicit one. For example, based on the scaling of fecundity per bout to reproductive mass per year for birds, we can predict that a cockatoo (*Cacatua ophthalmica*) with a fecundity of two eggs per bout (Myhrvold et al., 2015) produces a reproductive mass per year of *c*. 41 g (clade-specific conversion factors can be found in Supporting Information Table S8). Nonetheless, the conversions we provide here are applicable to only a few well-studied taxa, and even in these clades there is significant scatter. In particular, the low precision of predicting reproductive mass per year from reproductive traits without volumetric dimensions (i.e., fecundity per bout and fecundity per year) could lead to inaccurate estimates of reproductive output for some species. To illustrate, a turkey (*Meleagris ocellata*) with a fecundity of 12 eggs per bout (Myhrvold et al., 2015) is predicted to produce a reproductive mass per year of *c*. 87 g, but a single turkey egg weighs *c*. 90 g (Tserveni-Goussi & Fortomaris, 2011); consequently, the use of our conversion factors underestimates reproduction by around an order of magnitude for this species. Hence, our conversion factors are an improvement on using ambiguous measures, but their lack of precision makes them an imperfect solution. We encourage the exploration of reproductive trait relationships across additional animal classes and at lower taxonomic levels to determine how they compare to the few class-specific relationships we present here.

TABLE 2 Intercept (*a*) and slope (*b*) coefficients [±95% confidence intervals (CI)] of linear regression models between $\log_{10}$(ambiguous reproductive trait) and $\log_{10}$[reproductive mass per year (in grams per year)] for each animal class with ≥100 species. *Note*: Wald test statistics (*W*) assessed formally whether the intercept and slope coefficients were significantly different from zero and one, respectively ($p < .05$). The R^2 values provide an estimate of how well each ambiguous reproductive measure explains reproductive mass per year; values of one are perfect fits

Ambiguous reproductive trait	Intercept (*a*) Estimate (±CI)	Intercept (*a*) *W*	Intercept (*a*) *p*-value	Slope (*b*) Estimate (±CI)	Slope (*b*) *W*	Slope (*b*) *p*-value	R^2
Fecundity per bout (number of offspring per event)							
Aves	1.49 (±0.07)	44.8	<.001	0.42 (±0.11)	10.3	<.001	0.03
Mammalia	2.21 (±0.08)	56.6	<.001	−0.92 (±0.18)	21.4	<.001	0.06
Reptilia	−0.47 (±0.08)	11.1	<.001	2.12 (±0.11)	20.1	<.001	0.59
Fecundity per year (number of offspring per year)							
Aves	1.66 (±0.06)	51.2	<.001	0.08 (±0.09)	20.4	<.001	0.00
Mammalia	2.25 (±0.07)	60.8	<.001	−0.67 (±0.10)	31.9	<.001	0.09
Reptilia	−0.65 (±0.10)	12.7	<.001	1.84 (±0.11)	15.8	<.001	0.54
Reproductive mass per bout (g)							
Aves	0.27 (±0.02)	29.0	<.001	0.89 (±0.01)	19.9	<.001	0.94
Mammalia	0.36 (±0.02)	35.1	<.001	0.89 (±0.01)	22.3	<.001	0.89
Reptilia	0.23 (±0.02)	20.4	<.001	0.95 (±0.02)	6.29	<.001	0.93

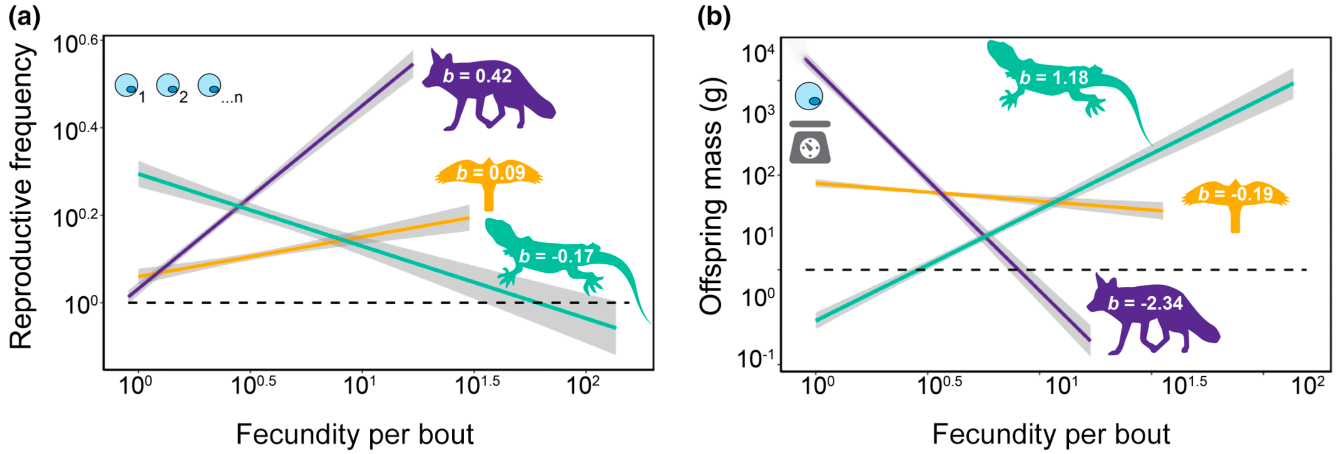


FIGURE 5 The relationship between fecundity per bout (as number of offspring per event) and: (a) reproductive frequency (as reproductive events per year) and (b) offspring mass (in grams) (shadings are ±95% confidence intervals). Scaling exponents of each vertebrate class are printed on the matching silhouette. The dashed line shows the null hypothesis for no covariance between ambiguous reproductive traits (i.e., $b = 0$). Silhouettes and colours are the same as in Figure 2

In comparison to other ambiguous measures, reproductive mass per bout was the most consistent predictor of reproductive mass per year. For well-studied vertebrates at least, offspring mass varies more among species than does frequency of reproductive bouts. As such, failure to incorporate offspring mass into calculations of reproductive output leads to greater misestimation than does exclusion of reproductive frequency, highlighting the importance of studies that seek to understand offspring-size variation more generally (Charnov & Ernest, 2006; Marshall et al., 2018; Neuheimer et al., 2015; Ronget et al., 2018). Given that coupling estimates of offspring mass to fecundity per bout (i.e., reproductive mass per bout) provides the least misleading proxy for reproductive mass per year, we suggest that future studies should prioritize the estimation of offspring mass ahead of reproductive frequency. Fortunately, weighing offspring is generally easier than estimating reproductive frequency; the former does not require monitoring reproduction over extended periods of time. We should note that further analyses are required to understand whether our results apply to under-represented taxonomic clades (e.g., insects and fishes). For these groups, reproductive frequency could vary as much as offspring size (Bakewell et al., 2020; Marshall et al., 2021).

Although the combination of estimates of offspring size and fecundity per bout greatly improves our capacity to predict reproductive mass per year, caution is still required. Offspring size metrics were inconsistent across databases and clades and were often estimated as length, despite mass being a more appropriate representation of total energy and reproductive investment provided per offspring (Barneche et al., 2018; Hatton et al., 2019; Pettersen et al., 2015). Although conversion of offspring length into mass is an option for some taxa (shown here in Figure 2, and elsewhere: Neuheimer et al., 2015), offspring length-to-mass conversions have similar risks to adult length-to-mass conversions; across large domain ranges, slight differences in length–mass scaling can yield serious misestimates (Feldman & Meiri, 2013; Froese & Pauly, 2021; Meiri, 2010; Neuheimer et al., 2015). Measurement of the mass of individual offspring can be challenging, especially for offspring that are relatively small (less than millimetre sized), and different metrics exist (e.g., wet weight, dry weight and ash-free weight) that reduce the compatibility of offspring size traits further. Despite its challenges, we advocate the measurement of offspring size directly as mass or, when this is not possible, the development of species-specific length-to-mass conversions, such that offspring sizes are compatible across taxa. Better still is to measure reproductive mass per bout directly, given that this measure was the best proxy for reproductive mass per year for several vertebrate clades. Fortunately, this can be particularly favourable when measuring the mass of a single offspring is difficult or impossible, such as when eggs are extremely small and difficult to isolate from the clutch (e.g., for species that lay gelatinous egg masses).

The compilation of reproductive traits has been a remarkable effort that has allowed researchers to access broad-scale questions (Fritz et al., 2009; Newbold et al., 2015; Tucker et al., 2018). Fortunately, reproductive trait databases do not appear to be limited by the inaccessibility issues that others have highlighted for databases of other traits (e.g., Hipsley & Sherratt, 2019; Salguero-Gómez et al., 2021; Savage & Vickers, 2009; Wicherts et al., 2006). We want to emphasize that our critique regarding the under-reporting of various reproductive measures is not a criticism of the original data collectors, nor those who assembled these databases. However, we must begin to acknowledge that our ability to compare reproductive output across taxa is currently limited by trait dimensionality (both volumetric and temporal), particularly given that such measures can have valuable applications for biological resource management (Marshall et al., 2021). For several well-studied vertebrate clades, we have provided relationships to convert temporally and volumetrically ambiguous measures to the most explicit measure of reproductive output (i.e., reproductive mass per year). For lesser-studied taxa, such as invertebrates, we think the hard work of collecting reproductive data is unavoidable, because other tactics to infer reproductive trait values, such as imputation, are unviable and prone to error when raw data are deficient (González-Suárez et al., 2012; Johnson et al., 2021). For invertebrates in particular, data imputation is impossible to estimate reproductive mass per year, because this key reproductive trait is missing entirely. To make comparisons across broader taxonomic clades, we advocate the collection of reproductive mass per year explicitly across more taxa. When this is not feasible, researchers should prioritize the measurement of reproductive mass per bout, because this trait was found to be the best proxy for reproductive mass per year. Reproductive trait compilations represent centuries of work, and we must continue to tie reproductive output to explicit temporal and volumetric dimensions to unlock their full potential to make meaningful comparisons of reproduction across animals.

ACKNOWLEDGMENTS

We would like to thank F. Chichorro, P. Ferns, N. Kolm, S. Téletchéa, F. Téletchéa and M. Valcu for kindly providing data via email. We also thank S. Browne, B. Comerford, G. Jarvis, E. Morris, M. Parascandalo, A. Potter and E. Richardson for advice on figures, and F. Schrodt, R. Salguero-Gómez and several anonymous reviewers for valuable feedback that improved this manuscript. We acknowledge support from Monash University and the Centre for Geometric Biology. Open access publishing facilitated by Monash University, as part of the Wiley - Monash University agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data are available in the online supplemental file All data are archived through Dryad <https://doi.org/10.5061/dryad.2547d7wts>.

ORCID

Samuel C. Ginther https://orcid.org/0000-0003-3263-8546
Hayley Cameron https://orcid.org/0000-0001-5004-6646
Craig R. White https://orcid.org/0000-0002-0200-2187
Dustin J. Marshall https://orcid.org/0000-0001-6651-6219

REFERENCES

Bakewell, A. T., Davis, K. E., Freckleton, R. P., Isaac, N. J. B., & Mayhew, P. J. (2020). Comparing life histories across taxonomic groups in multiple dimensions: How mammal-like are insects? *The American Naturalist*, *195*, 70–81.

Barneche, D. R., Robertson, D. R., White, C. R., & Marshall, D. J. (2018). Fish reproductive-energy output increases disproportionately with body size. *Science*, *360*, 642–645.

Blueweiss, L., Fox, H., Kudzma, V., Nakashima, D., Peters, R., & Sams, S. (1978). Relationships between body size and some life history parameters. *Oecologia*, *37*, 257–272.

Burger, J. R., Hou, C., & Brown, J. H. (2019). Toward a metabolic theory of life history. *Proceedings of the National Academy of Sciences*, *116*, 26653–26661.

Chamberlain, S. A., & Szöcs, E. (2013). Taxize: Taxonomic search and retrieval in R [version 2; peer review: 3 approved]. *F1000 Research*, *2*, 1–30.

Charnov, E. L., & Ernest, S. K. M. (2006). The offspring-size/clutch-size trade-off in mammals. *American Naturalist*, *167*, 578–582.

Dziock, F., Gerisch, M., Siegert, M., Hering, I., Scholz, M., & Ernst, R. (2011). Reproducing or dispersing? Using trait based habitat templet models to analyse orthoptera response to flooding and land use. *Agriculture, Ecosystems & Environment*, *145*, 85–94.

Ernest, S. K. M., Enquist, B. J., Brown, J. H., Charnov, E. L., Gillooly, J. F., Savage, V. M., White, E. P., Smith, F. A., Hadly, E. A., Haskell,

J. P., Lyons, S. K., Maurer, B. A., Niklas, K. J., & Tiffney, B. (2003). Thermodynamic and metabolic effects on the scaling of production and population energy use: Thermodynamic and metabloic effects. *Ecology Letters*, *6*, 990–995.

Faulwetter, S., Markantonatou, V., Pavloudi, C., Papageorgiou, N., Keklikoglou, K., Chatzinikolaou, E., Pafilis, E., Chatzigeorgiou, G., Vasileiadou, K., Dailianis, T., Fanini, L., Koulouri, P., & Arvanitidis, C. (2014). Polytraits: A database on biological traits of marine polychaetes. *Biodiversity Data Journal*, *2*, e1024.

Feldman, A., & Meiri, S. (2013). Length-mass allometry in snakes: Snake length-mass allometry. *Biological Journal of the Linnean Society*, *108*, 161–172.

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed). Thousand Oaks, CA: Sage. http://socialsciences.mcmaster.ca/jfox/Books/Companion

Fritz, S. A., Bininda-Emonds, O. R. P., & Purvis, A. (2009). Geographical variation in predictors of mammalian extinction risk: Big is bad, but only in the tropics. *Ecology Letters*, *12*, 538–549.

Froese, R. & Pauly, D. (Eds.) (2021). *Fishbase*. World Wide Web electronic publication www.fishbase.org

González-Suárez, M., Lucas, P. M., & Revilla, E. (2012). Biases in comparative analyses of extinction risk: Mind the gap. *Journal of Animal Ecology*, *81*, 1211–1222.

Hatton, I. A., Dobson, A. P., Storch, D., Galbraith, E. D., & Loreau, M. (2019). Linking scaling laws across eukaryotes. *Proceedings of the National Academy of Sciences*, *116*, 21616–21622.

Healy, K., Ezard, T. H. G., Jones, O. R., Salguero-Gómez, R., & Buckley, Y. M. (2019). Animal life history is shaped by the pace of life and the distribution of age-specific mortality and reproduction. *Nature Ecology & Evolution*, *3*, 1217–1224.

Hipsley, C. A., & Sherratt, E. (2019). Psychology, not technology, is our biggest challenge to open digital morphology data. *Scientific Data*, *6*, 41.

James, K. L., Randall, N. P., & Haddaway, N. R. (2016). A methodology for systematic mapping in environmental sciences. *Environmental Evidence*, *5*, 7.

Jelbert, K., Buss, D., McDonald, J., Townley, S., Franco, M., Stott, I., Jones, O., Salguero-Gómez, R., Buckley, Y., Knight, T., Silk, M., Sargent, F., Rolph, S., Wilson, P., & Hodgson, D. (2019). Demographic amplification is a predictor of invasiveness among plants. *Nature Communications*, *10*, 5602.

Jervis, M., & Ferns, P. (2011). Towards a general perspective on LIFE-history evolution and diversification in PARASITOID wasps: PARASITOID WASP LIFE-HISTORY PREDICTORS. *Biological Journal of the Linnean Society*, *104*, 443–461.

Johnson, T. F., Isaac, N. J. B., Paviolo, A., & González-Suárez, M. (2021). Handling missing values in trait data. *Global Ecology and Biogeography*, *30*, 51–62.

Jones, K. E., Bielby, J., Cardillo, M., Fritz, S. A., O'Dell, J., Orme, C. D. L., Safi, K., Sechrest, W., Boakes, E. H., Carbone, C., Connolly, C., Cutts, M. J., Foster, J. K., Grenyer, R., Habib, M., Plaster, C. A., Price, S. A., Rigby, E. A., Rist, J., … Purvis, A. (2009). PanTHERIA: A species-level database of life history, ecology, and geography of extant and recently extinct mammals: *Ecological archives* E090-184. *Ecology*, *90*, 2648.

Kokko, H. (2021). The stagnation paradox: The ever-improving but (more or less) stationary population fitness. *Proceedings of the Royal Society B: Biological Sciences*, *288*, 20212145.

Lack, D. (1947). The significance of clutch-size. *Ibis*, *89*, 302–352.

Marshall, D. J., Barneche, D. R., & White, C. R. (2021). How does spawning frequency scale with body size in marine fishes? *Fish and Fisheries*, *23*, 316–323.

Marshall, D. J., Pettersen, A. K., & Cameron, H. (2018). A global synthesis of offspring size variation, its eco-evolutionary causes and consequences. *Functional Ecology*, *32*, 1436–1446.

Meiri, S. (2010). Length–weight allometries in lizards. *Journal of Zoology*, *281*, 218–226.

Meiri, S., Brown, J. H., & Sibly, R. M. (2012). The ecology of lizard reproductive output. *Global Ecology and Biogeography*, *21*, 592–602.

Mills, J. A., Teplitsky, C., Arroyo, B., Charmantier, A., Becker, P. H., Birkhead, T. R., Bize, P., Blumstein, D. T., Bonenfant, C., Boutin, S., Bushuev, A., Cam, E., Cockburn, A., Côté, S. D., Coulson, J. C., Daunt, F., Dingemanse, N. J., Doligez, B., Drummond, H., … Zedrosser, A. (2015). Archiving primary data: Solutions for long-term studies. *Trends in Ecology & Evolution*, *30*, 581–589.

Myhrvold, N. P., Baldridge, E., Chan, B., Sivam, D., Freeman, D. L., & Ernest, S. K. M. (2015). An amniote life-history database to perform comparative analyses with birds, mammals, and reptiles: *Ecological archives* E096-269. *Ecology*, *96*, 3109.

Neuheimer, A. B., Hartvig, M., Heuschele, J., Hylander, S., Kiørboe, T., Olsson, K. H., Sainmont, J., & Andersen, K. H. (2015). Adult and offspring size in the ocean over 17 orders of magnitude follows two life history strategies. *Ecology*, *96*, 3303–3311.

Newbold, T., Hudson, L. N., Hill, S. L. L., Contu, S., Lysenko, I., Senior, R. A., Börger, L., Bennett, D. J., Choimes, A., Collen, B., Day, J., De Palma, A., Díaz, S., Echeverria-Londoño, S., Edgar, M. J., Feldman, A., Garon, M., Harrison, M. L. K., Alhusseini, T., … Purvis, A. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*, *520*, 45–50.

Nørgaard, L. S., Álvarez-Noriega, M., McGraw, E., White, C. R., & Marshall, D. J. (2021). Predicting the response of disease vectors to global change: The importance of allometric scaling. *Global Change Biology*, *28*, 390–402.

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, *5*, 210.

Pettersen, A. K., White, C. R., & Marshall, D. J. (2015). Why does offspring size affect performance? Integrating metabolic scaling with life-history theory. *Proceedings of the Royal Society B: Biological Sciences*, *282*, 20151946.

Ronget, V., Gaillard, J.-M., Coulson, T., Garratt, M., Gueyffier, F., Lega, J.-C., & Lemaître, J.-F. (2018). Causes and consequences of variation in offspring body mass: Meta-analyses in birds and mammals. *Biological Reviews*, *93*, 1–27.

Salguero-Gómez, R., Jackson, J., & Gascoigne, S. J. L. (2021). Four key challenges in the open-data revolution. *Journal of Animal Ecology*, *90*, 2000–2004.

Salguero-Gómez, R., Jones, O. R., Archer, C. R., Bein, C., Buhr, H., Farack, C., Gottschalk, F., Hartmann, A., Henning, A., Hoppe, G., Römer, G., Ruoff, T., Sommer, V., Wille, J., Voigt, J., Zeh, S., Vieregg, D., Buckley, Y. M., Che-Castaldo, J., … Vaupel, J. W. (2016). COMADRE: A global data base of animal demography. *Journal of Animal Ecology*, *85*, 371–384.

Sarremejane, R., Cid, N., Stubbington, R., Datry, T., Alp, M., Cañedo-Argüelles, M., Cordero-Rivera, A., Csabai, Z., Gutiérrez-Cánovas, C., Heino, J., Forcellini, M., Millán, A., Paillex, A., Pařil, P., Polášek, M., Tierno de Figueroa, J. M., Usseglio-Polatera, P., Zamora-Muñoz, C., & Bonada, N. (2020). DISPERSE, a trait database to assess the dispersal potential of European aquatic macroinvertebrates. *Scientific Data*, *7*, 386.

Savage, C. J., & Vickers, A. J. (2009). Empirical study of data sharing by authors publishing in PLoS journals. *PLoS One*, *4*, e7078.

Savage, V. M., Gillooly, J. F., Brown, J. H., West, G. B., & Charnov, E. L. (2004). Effects of body size and temperature on population growth. *The American Naturalist*, *163*, 429–441.

Thorson, G. (1950). Reproductive and larval ecology of marine bottom invertebrates. *Biological Reviews*, *25*, 1–45.

Tserveni-Goussi, A., & Fortomaris, P. (2011). Production and quality of quail, pheasant, goose and Turkey eggs for uses other than human consumption. In Y. Nys, M. Bain, & F. Van Immerseel (Eds.), *Improving the safety and quality of eggs and egg products* (pp. 509–537). Elsevier.

Tucker, M. A., Böhning-Gaese, K., Fagan, W. F., Fryxell, J. M., Van Moorter, B., Alberts, S. C., Ali, A. H., Allen, A. M., Attias, N., Avgar, T., Bartlam-Brooks, H., Bayarbaatar, B., Belant, J. L., Bertassoni, A., Beyer, D., Bidner, L., van Beest, F. M., Blake, S., Blaum, N., … Mueller, T. (2018). Moving in the Anthropocene: Global reductions in terrestrial mammalian movements. *Science*, *359*, 466–469.

White, C. R., Marshall, D. J., Alton, L. A., Arnold, P. A., Beaman, J. E., Bywater, C. L., Condon, C., Crispin, T. S., Janetzki, A., Pirtle, E., Winwood-Smith, H. S., Angilletta, M. J., Chenoweth, S. F., Franklin, C. E., Halsey, L. G., Kearney, M. R., Portugal, S. J., & Ortiz-Barrientos, D. (2019). The origin and maintenance of metabolic allometry in animals. *Nature Ecology & Evolution*, *3*, 598–603.

White, C. R., Marshall, D. J., Chown, S. L., Clusella-Trullas, S., Portugal, S. J., Franklin, C. E., & Seebacher, F. (2021). Geographical bias in physiological data limits predictions of global change impacts. *Functional Ecology*, *35*, 1572–1578.

Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *American Psychologist*, *61*, 726–728.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*, 1686.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018.

BIOSKETCH

Samuel C. Ginther studies life-history evolution through comparative and experimental approaches.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ginther, S. C., Cameron, H., White, C. R., & Marshall, D. J. (2022). Avoiding growing pains in reproductive trait databases: the curse of dimensionality. *Global Ecology and Biogeography*, *00*, 1–13. https://doi.org/10.1111/geb.13589

APPENDIX 1

Data Sources

Beukhof, E., Frelat, R., Pecuchet, L., Maureaud, A., Dencker, T. S., Sólmundsson, J., Punzón, A., Primicerio, R., Hidalgo, M., Möllmann, C., & Lindegren, M. (2019). Marine fish traits follow fast-slow continuum across oceans. *Scientific Reports*, **9**, 17878.

Bird, J. P., Martin, R., Akçakaya, H. R., Gilroy, J., Burfield, I. J., Garnett, S. T., Symes, A., Taylor, J., Şekercioğlu, Ç. H., & Butchart, S. H. M. (2020). Generation lengths of the world's birds and their implications for extinction risk. *Conservation Biology*, **34**, 1252–1261.

Brun, P., Payne, M. R., & Kiørboe, T. (2017). A trait database for marine copepods. *Earth System Science Data*, **9**, 99–113.

Cofre, H. L., Böhning-Gaese, K., & Marquet, P. A. (2007). Rarity in Chilean forest birds: which ecological and life-history traits matter? *Diversity and Distributions*, **13**, 203–212.

Cooper, N., Bielby, J., Thomas, G. H., & Purvis, A. (2008). Macroecology and extinction risk correlates of frogs. *Global Ecology and Biogeography*, **17**, 211–221.

De Magalhães, J. P., & Costa, J. (2009). A database of vertebrate longevity records and their relation to other life-history traits. *Journal of Evolutionary Biology*, **22**, 1770–1774.

Ellers, J., Berg, M. P., Dias, A. T. C., Fontana, S., Ooms, A., & Moretti, M. (2018). Diversity in form and function: Vertical distribution of soil fauna mediates multidimensional trait variation. *Journal of Animal Ecology*, **87**, 933–944.

Etard, A., Morrill, S., & Newbold, T. (2020). Global gaps in trait data for terrestrial vertebrates. *Global Ecology and Biogeography*, **29**, 2143–2158.

Freshwater, C., Ghalambor, C. K., & Martin, P. R. (2014). Repeated patterns of trait divergence between closely related dominant and subordinate bird species. *Ecology*, **95**, 2334–2345.

Frimpong, E. A., & Angermeier, P. L. (2009). Fish Traits: A database of ecological and life-history traits of freshwater fishes of the United States. *Fisheries*, **34**, 487–495.

Froese, R., & Pauly, D. (2021). Fishbase. www.fishbase.org

Harabis, F., & Hronkova, J. (2020). European database of the life-history, morphological and habitat characteristics of dragonflies (Odonata). *European Journal of Entomology*, **117**, 302–308.

Jervis, M., & Ferns, P. (2011). Toward a general perspective on life-history evolution and diversification in parasitoid wasps. *Biological Journal of the Linnean Society*, **104**, 443–461.

Kindsvater, H. K., Braun, D. C., Otto, S. P., & Reynolds, J. D. (2016). Costs of reproduction can explain the correlated evolution of semelparity and egg size: theory and a test with salmon. *Ecology Letters*, **19**, 687–696.

Lecocq, T., Benard, A., Pasquet, A., Nahon, S., Ducret, A., Dupont-Marin, K., Lang, I., & Thomas, M. (2019). TOFF, a database of traits of fish to promote advances in fish aquaculture. *Scientific Data*, **6**, 301.

Liedtke, H. C., Müller, H., Hafner, J., Nagel, P., & Loader, S. P. (2014). Interspecific patterns for egg and clutch sizes of African Bufonidae (Amphibia: Anura). *Zoologischer Anzeiger - A Journal of Comparative Zoology*, **253**, 309–315.

Lotze, H. K., Flemming, J. M., & Magera, A. M. (2017). Critical factors for the recovery of marine mammals. *Conservation Biology*, **31**, 1301–1311.

Madin, J. S., Anderson, K. D., Andreasen, M. H., Bridge, T. C. L., Cairns, S. D., Connolly, S. R., Darling, E. S., Diaz, M., Falster, D. S., Franklin, E. C., Gates, R. D., Harmer, A. M. T., Hoogenboom, M. O., Huang, D., Keith, S. A., Kosnik, M. A., Kuo, C.-Y., Lough, J. M., Lovelock, C. E., …, Baird, A. H. (2016). The Coral Trait Database, a curated database of trait information for coral species from the global oceans. *Scientific Data*, **3**, 160017.

Meiri, S., Brown, J. H., & Sibly, R. M. (2012). The ecology of lizard reproductive output. *Global Ecology and Biogeography*, **21**, 592–602.

Mesquita, D. O., Costa, G. C., Colli, G. R., Costa, T. B., Shepard, D. B., Vitt, L. J., & Pianka, E. R. (2016). Life-history patterns of lizards of the world. *The American Naturalist*, **187**, 689–705.

Minias, P., & Włodarczyk, R. (2020). Avian developmental rates are constrained by latitude and migratoriness – A global analysis. *Journal of Biogeography*, **47**, 2156–2167.

Moore, A. P., Galic, N., Brain, R. A., Hornbach, D. J., & Forbes, V. E. (2021a). Validation of freshwater mussel life-history strategies: A database and multivariate analysis of freshwater mussel life-history traits. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **31**, 3386–3402.

Moore, C. E., Helmann, J. S., Chen, Y., St. Amour, S. M., Hallmark, M. A., Hughes, L. E., Wax, N., & Mims, M. C. (2021b). Anuran traits of the United States (ATraiU): A database for anuran traits-based conservation, management, and research. *Ecology*, **102**.

Morrow, C. B., Ernest, S. K. M., & Kerkhoff, A. J. Macroevolution of dimensionless life-history metrics in tetrapods. *Proceedings of the Royal Society B: Biological Sciences*, **288**, 20210200.

Oliveira, B. F., São-Pedro, V. A., Santos-Barrera, G., Penone, C., & Costa, G. C. (2017). AmphiBIO, a global database for amphibian ecological traits. *Scientific Data*, **4**, 170123.

Pekár, S., Wolff, J. O., Černecká, Ľ., Birkhofer, K., Mammola, S., Lowe, E. C., Fukushima, C. S., Herberstein, M. E., Kučera, A., Buzatto, B. A., Djoudi, E. A., Domenech, M., Enciso, A. V., Piñanez Espejo, Y. M. G., Febles, S., García, L. F., Gonçalves-Souza, T., Isaia, M., Lafage, D., …, Cardoso, P. (2021). The World Spider Trait database: a centralised global open repository for curated data on spider traits. *Database*, **2021**, baab064.

Pianka, E. R., Vitt, L. J., Pelegrin, N., Fitzgerald, D. B., & Winemiller, K. O. (2017). Toward a periodic table of niches, or exploring the lizard niche hypervolume. *The American Naturalist*, **190**, 601–616.

Razafindratsima, O. H., Yacoby, Y., & Park, D. S. (2018). MADA: Malagasy Animal trait Data Archive. *Ecology*, **99**, 990–990.

Sibly, R. M., Witt, C. C., Wright, N. A., Venditti, C., Jetz, W., & Brown, J. H. (2012). Energetics, lifestyle, and reproduction in birds. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 10937–10941.

Soria, C. D., Pacifici, M., Di Marco, M., Stephen, S. M., & Rondinini, C. (2021). COMBINE: a coalesced mammal database of intrinsic and extrinsic traits. *Ecology*, **102**.

Storchová, L., & Hořák, D. (2018). Life-history characteristics of European birds. *Global Ecology and Biogeography*, **27**, 400–406.

Teletchea, S., & Teletchea, F. STOREFISH 2.0: A database on the reproductive strategies of teleost fishes. *Database*, **2020**, baaa095.

Trakimas, G., Whittaker, R. J., & Borregaard, M. K. (2016). Do biological traits drive geographical patterns in European amphibians? *Global Ecology and Biogeography*, **25**, 1228–1238.

Trochet, A., Moulherat, S., Calvez, O., Stevens, V., Clobert, J., & Schmeller, D. (2014). A database of life-history traits of European amphibians. *Biodiversity Data Journal*, **2**, e4123.

Tsai, P., Ko, C., Hsieh, C., Su, Y., Lu, Y., Lin, R., & Tuanmu, M. (2020). A trait dataset for Taiwan's breeding birds. *Biodiversity Data Journal*, **8**, e49735.

Vanadzina, K., Phillips, A., Martins, B., Laland, K. N., Webster, M. M., & Sheard, C. (2021). Ecological and behavioural drivers of offspring size in marine teleost fishes. *Global Ecology and Biogeography*, **30**, 2407–2419.

Van Wilgen, N. J., & Richardson, D. M. (2012). The roles of climate, phylogenetic relatedness, introduction effort, and reproductive traits in the establishment of non-native reptiles and amphibians. *Conservation Biology*, **26**, 267–277.