New horizons for comparative studies and meta-analyses

Patrice Pottier1*, Daniel W.A. Noble2+, Frank Seebacher3+, Nicholas C. Wu4+, Samantha Burke1,
Malgorzata Lagisz1,5, Lisa E. Schwanz1, Szymon M. Drobnia1,6#, Shinichi Nakagawa1,5#

1 Evolution & Ecology Centre, School of Biological, Earth and Environmental Sciences, UNSW, Sydney, Australia.
2 Division of Ecology and Evolution, Research School of Biology, College of Science, The Australian National University, Canberra, Australian Capital Territory, Australia.
3 School of Life and Environmental Sciences, University of Sydney, Sydney, New South Wales 2006, Australia.
4 Hawkesbury Institute for the Environment, Western Sydney University, NSW 2753, Australia
5 Theoretical Sciences Visiting Program, Okinawa Institute of Science and Technology Graduate University, Onna, 904-0495, Japan.
6 Institute of Environmental Sciences, Jagiellonian University, Krakow, Poland.

*Corresponding author: Pottier, P. (p.pottier@unsw.edu.au); Twitter: @PatriceEcoEvo
#These authors supervised this work equally
+These authors contributed equally

ORCID

Patrice Pottier https://orcid.org/0000-0003-2106-6597
Daniel W.A. Noble https://orcid.org/0000-0001-9460-8743
Frank Seebacher https://orcid.org/0000-0002-2281-9311
Nicholas C. Wu https://orcid.org/0000-0002-7130-1279
Samantha Burke https://orcid.org/0000-0001-6902-974X
Malgorzata Lagisz https://orcid.org/0000-0002-3993-6127
Lisa E. Schwanz https://orcid.org/0000-0001-5864-7112
Szymon M. Drobnia https://orcid.org/0000-0001-8101-6247
Shinichi Nakagawa https://orcid.org/0000-0002-7765-5182
**Highlights**

Meta-analyses are often used in ecology but originated in the medical and social sciences, whereas phylogenetic comparative analyses stemmed from evolutionary biology.

We show that these two methods can be mathematically equivalent although their current use in ecology and evolution has different strengths and limitations.

We advocate that integrating their strengths will improve the accuracy, robustness, and transparency of ecological and evolutionary syntheses, resolving issues such as missing data and publication bias and opening new avenues of research.

We highlight future opportunities, such as exploring complex (non-linear) trends, testing hypotheses across multiple scales and levels of organization, and calling for big-team science collaboration to conduct ‘prospective’ and ‘living’ comparative and meta-analyses.
Abstract

Comparative analyses and meta-analyses are key tools to elucidate broad biological principles, yet the two approaches often appear different in purpose. We propose an integrated approach that can generate deeper insights into eco-evolutionary processes. Marrying comparative and meta-analytic approaches will allow for 1) a more accurate investigation of drivers of biological variation; 2) a greater ability to account for sources of non-independence in experimental data; 3) more effective control of publication bias; and 4) improved transparency and reproducibility. Stronger integration of meta-analytic and comparative studies can also broaden the scope from species-centric investigations to community-level responses and function-valued traits (e.g., reaction norms). We illuminate commonalities, differences, and the transformative potential of combining these methodologies for advancing ecology and evolutionary biology.

Keywords: multilevel modelling, multivariate analysis, phylogenetic generalized linear mixed model, sampling variance, phylogenetic signal
History and purpose of comparative studies and meta-analyses

Differences between species have inspired biological research from antiquity to recent analyses of biodiversity [1,2]. Darwin contributed to the foundation of modern biology through his comparisons between species and explanations of the origins of species divergence [3]. Comparisons between species became a focus for uncovering biological patterns and processes. Well known examples include Bergman's rule stating that animal size increases with increasing latitude [4], or the mouse-to-elephant curve of metabolic rates as a function of animal mass [5]. The common approach of these and other comparative studies was to analyse species-level traits to uncover evolutionary principles that explain trait variation at the tips of phylogenies (i.e., present day).

Increasing sophistication of phylogenetic comparative methods and the construction of phylogenetic trees in the 20th century led to a greater appreciation of the ancestral connectivity between species in deep time [6]. It became clear that species are connected to varying degrees by their common ancestry, which can potentially confound (uncorrected) species comparisons. The comparative method now explicitly addresses the problem of phylogenetic non-independence (see Glossary), using statistical approaches to account for species relatedness in phylogenetic comparative analyses (see Glossary) [7–11]. Modern phylogenetic comparative analyses thereby incorporate ancestral state/trait reconstruction (see Glossary) and even time calibrations (see Glossary) of divergence between lineages with the goal to understand the evolutionary processes that gave rise to trait differences [12–14].

Similar to comparative analyses, biological meta-analyses (see Glossary) typically compare traits across taxa. Phylogenetic correction is therefore essential here too. However, unlike comparative analyses, meta-analyses typically do not ask explicit questions about evolutionary processes and are primarily focussed on present-day phenotypes [15]. Meta-
Analysis was developed in medicine and social sciences in parallel with evolutionary comparative analyses [16,17]. The purpose of early meta-analyses was to synthesize results across similar studies to detect the overall magnitude of treatment effects in a single species (humans). Eventually, meta-analyses were extended to ecology and evolution to integrate data from multiple species [18]. Their purpose shifted from simple quantification of effects to establishing the state-of-knowledge in a field and reassessing established hypotheses with mixed empirical support (e.g., [19,20]). Contrary to traditional comparative analyses, biological meta-analyses purposefully use heterogeneous datasets to estimate effect variation among multiple sources (e.g., different populations, species, geographical areas), while explicitly accounting for the variation due to sampling effort [17]. Meta-analyses also provide ways to test for publication biases. Such analyses not only assess the reliability of research findings but can also illuminate social dimensions in the research and publishing process [21].

Here, we argue that integrating comparative studies with meta-analyses will significantly advance progress in ecology and evolution. Attempts to merge these approaches have been proposed earlier [22–24], yet the two approaches are still rarely considered as complementary. Even though comparative analyses and meta-analyses appear different in their purpose, both have similar, if not identical, mathematical foundations (Box 1) and often address similar questions. It is therefore relatively straightforward to foster a greater synergy between these approaches. Here, we unify the two approaches conceptually and practically to analyse biological variation beyond species means (Figure 1). We also demonstrate how the analysis of function-valued traits and community-level patterns can generate new insights into eco-evolutionary processes (Figure 2). We show that the integration of comparative studies and meta-analyses provides better ways to explore biological variation and leads to greater transparency and reproducibility (see Glossary).
A unified approach for comparative studies

A multilevel framework to decompose biological and methodological variation

A key similarity between comparative and meta-analyses is the possibility to use a multilevel framework (see Glossary), which is the gold standard for meta-analyses [23] and phylogenetic generalized linear mixed models (PGLMM; see Glossary) [25–27]. In fact, PGLMM that incorporate sampling variance (see Glossary) are statistically identical to phylogenetic multilevel meta-analyses [23,24] (Box 1). However, comparative analyses often do not use a multilevel framework, focusing primarily on species mean-trait values, which are sometimes derived from few specimens [28–30]. This approach neglects within-species variation, which can bias results [29,30]. A multilevel framework, on the other hand, allows the partition of variance into different categories based on the natural hierarchical structure of the dataset (e.g., multiple species, populations, studies), thereby ultimately identifying biological and methodological drivers of the observed patterns. In fact, quantifying among-species variation can only be achieved when the importance of other contributors of phenotypic variation are considered explicitly [31]. For instance, the accuracy of estimating a phylogenetic signal (see Glossary) depends on quantifying the relative contribution of other sources of variation, such as epigenetics, species ecology and study-specific effects, because otherwise these sources of variance may be confounded with phylogenetic variance [32]. A multilevel framework also addresses complex issues of biological and methodological non-independence (see Glossary), allowing one to synthesize data beyond species means by leveraging complex datasets [33,34]. Accounting for species phylogenetic non-independence also quantifies variation due to shared ancestry. However, the depth of phylogenetic methods used in comparative analyses surpasses those generally employed in meta-analyses. Meta-analyses could therefore benefit from using more sophisticated phylogenetic methods such as ancestral trait reconstructions, and using different models of evolution (e.g., Brownian vs. Brownian motion).
Implementing these methods can help decipher the evolutionary processes that led to present-day phenotypes - sparking new hypotheses and pioneering novel research avenues.

**Considering sampling variance to improve precision and account for publication bias**

A key distinction between traditional comparative studies and meta-analyses is that meta-analyses have greater accuracy in estimating biological variation [35]. Indeed, meta-analyses give less weight to effect sizes (see Glossary) with higher sampling variance (less precision). We argue that comparative analyses would achieve more precise conclusions by removing sampling variance from the total variance. When not accounted for, sampling variation affects how variance is estimated across the investigated sources (e.g., within-species, among-species, among-population variation, phylogenetic signal). Taxonomic chauvinism (see Glossary) [36] makes this issue particularly important for comparative studies. Some species are better studied than others, and the likelihood of detecting false or imprecise patterns is particularly high when species-level data are derived from few specimens (i.e., when the sampling variance is large) [37,38]. By extension, accounting for differences in sampling variance provides the opportunity to assess publication bias (see Glossary), a critical aspect that has been largely overlooked in traditional comparative analyses and that can greatly affect conclusions [39]. Investigating publication bias can reveal societal pressures obstructing the publication of relevant data. For instance, the “file-drawer problem” [40] is a type of research bias where studies with non-significant results often go unpublished. Identifying such biases can help more accurate interpretation of the available literature. Furthermore, recent tools can not only detect but also correct for publication bias, enabling interpretation of potentially unbiased estimates [41].
Integrating different effect size measures for greater flexibility

Meta-analyses and comparative analyses also differ in the effect measures (see Glossary) used. Effect measures can be categorized into three distinct types: 1) single-group measures (e.g., species trait mean, rate, proportion); 2) comparative measures for two groups (e.g., standardized mean difference [42], log response ratio [43], log variability ratio [44], odds ratio [45]); and 3) association measures between two variables (e.g., correlation coefficients, standardized slopes [46]). Comparative analyses primarily use the first type while meta-analyses frequently make use of the latter two [47]. The advantage of using single-group measures is that results are easy to interpret, but single-group measures are limited to one type of response variable and unit. On the other hand, the latter two effect measures are not as straightforward to interpret but are standardized and can be compared across traits measured in different units. We argue that comparative and meta-analyses can both make use of all three types of effect measures. Common meta-analytic models are not restricted to traditional effect sizes and can make use of single-group measures (e.g., trait means). Similarly, phylogenetic comparative analyses may use association or comparative measures when these effect measures are more readily available, or when analysing traits measured in different units.

Reporting guidelines to promote transparency and reproducibility

Another key distinction between comparative studies and meta-analyses pertains to the methods and reporting used to ensure reproducibility. The historical connectivity of meta-analysis with medicine and social sciences has generated guidelines and recommendations to ensure transparency and reproducibility [48,49]. Adopting reporting standards aids in study design, ensures the inclusion of important methodological details, and ultimately elevates the reliability of research. In ecology and evolutionary biology, established guidelines for
systematic reviews and meta-analyses, such as PRISMA-EcoEvo [50], can be adopted with little adjustments. However, there is a need to conceptualize reporting guidelines tailored specifically for phylogenetic comparative studies. Comparative studies may also benefit from using best practices developed for meta-analyses, as both approaches share similar tools and objectives, as we have argued so far. Ideally, literature-based data syntheses should stem from a systematic review (see Glossary) with fully documented and reproducible procedures [51]. This practice simplifies updates and replications which helps build more trust in evidence [52]. However, this is not always feasible, particularly because most data used in comparative studies are taken from data compendia [53]. Building upon previous data compilations to incorporate important information (i.e., data provenance, metadata, sampling variance) could elevate the robustness of future studies and broaden the applicability of data compendia for comparative- and meta-analyses.

New opportunities and directions for comparative studies

Community-level responses and eco-evolutionary patterns

Community- and ecosystem-level analyses are routinely used in plant and community ecology. Yet, use of community-level responses is limited in phylogenetic comparative analyses. This is perhaps because phylogenetic components cannot be modelled at the community level, though phylogeny can still be modelled within each community. For instance, Markovski and colleagues [54] recently investigated global variation in the relationship between population size and sexual dimorphism. The authors quantified community-level standardized slopes where a positive relationship between sexual dimorphism and population size is predicted if sexual selection promotes viability via good genes [55]. The study yielded 2,592 slopes and error variances from phylogenetically controlled analyses in each community (grid cell), and then estimated overall slopes for both
resident and migratory species. Notably, this study controlled for spatial dependence (see Glossary) across grid cells and decomposed sources of variance. The researchers found the expected benefits of sexual selection in resident species, but not in migratory. This is an elegant example of how marrying phylogenetic comparative analysis with spatially-controlled meta-analysis can lead to significant new insights. At the community level, species assemblages vary greatly, which provides interesting insights into ecological (e.g. spatial variation) and evolutionary (e.g., phylogenetic signal) processes governing biological variation at different scales. In addition, studying community-level patterns allow to capture the influence of species interactions that may be missed in individual- or species-level analyses. Such interactions (e.g., predator-prey, competition) are integral parts of an ecosystem and profoundly influence species traits within each community. Community-level analyses can also illuminate variability in species’ responses to environmental change across communities. This variability can be key in identifying species and communities more sensitive to disturbance.

Function-valued traits and multivariate meta-analyses

Function-valued traits are organismal responses to continuous variables such as temperature, pH, or age [56,57]. Examples include performance curves, growth trajectories, reflectance spectra, or sonograms. Function-valued traits can perhaps be seen as a new type of effect measure that combines a set of parameters. Complex function-valued traits are better depicted by curves than lines, and such traits can be summarized in various descriptors (parameters) of the curve (e.g., intercept, slope, peak, asymptote). The best way to model multi-parameter measures (i.e., function-valued traits) is with phylogenetic multivariate meta-analysis (see Glossary) (PMMA; [23]), which can estimate not only correlations between parameters of function-valued traits, but also accounts for uncertainty in measurements.
Notably, PMMA is an extension of phylogenetic comparative multivariate analysis, which is used to examine, for example, morphometric data where ‘landmarks’ describe complex traits such as body shapes [58–60].

Trait correlations are important to consider in function-valued traits, as one parameter (e.g., intercept) may constrain another parameter (e.g., slope). Multivariate models can quantify correlations between parameters of function-valued traits, which can highlight co-evolution and trade-offs among parameters [61]. For example, Pettersen and colleagues [62] collected data on temperature-dependent hatching success and estimated embryo optimal temperature (from a nonlinear function-valued trait). The researchers collected preferred body temperature of gravid and non-gravid females for >120 squamate species. Using multivariate models that accounted for sampling error, the authors estimated the phylogenetic relationships among these variables to understand how conflicts between embryo optimal temperatures and female preferred body temperatures are alleviated when gravid. The study shows that behavioural adjustments by gravid females can circumvent different thermal optima for embryos and mothers, and may help pave the way in explaining why viviparity evolves so regularly (>115 times) in squamates.

Another major benefit of using multivariate models for the analysis of function-valued traits is their potential to improve precision. When multiple parameters of the function-valued traits are correlated, precision around parameters can be improved by explicitly accounting for the covariance among variables (“borrowing of strength” [63]). Using multivariate models also means that some of the parameters of function-valued traits can be missing as long as not all trait values are missing for each species [64,65]. Indeed, phylogenetic multivariate models and related techniques can impute missing data, as missing trait values are inferred from the available parameters. Data imputation will extend not only the number of traits but also the number of species that can be examined, although the
effective use of imputation remains uncommon [62,66]. Taken together, the analysis of
function-valued traits could dramatically increase the number of biological questions that can
be asked, and better capture the intricate shape of biological responses.

Towards next-generation comparative studies via open synthesis communities

Community-level and function-valued analyses may require much larger datasets than
ordinary analyses. This calls for scientific “community-level” collaboration [67]. Movements
towards big-team science are already happening in the form of global research networks and
globally distributed experiments such as SPI-Birds [68], the Global Urban Evolution Project
[69], the Nutrient Network [70], and large-scale collaborative databases such as BioTIME
[71] or PREDICTS [72].

A rapid spread of global research networks provides a foundation for “open synthesis
communities”, where scientists with similar interests can plan and conduct comparative and
meta-analyses together with research synthesis specialists (e.g., librarians and information
scientists [73]). Simultaneously, such communities increasingly adhere to the principles of
open science, embracing open participation, materials, data, and code [74]. An open synthesis
community can carry out prospective meta-analyses as well as living/dynamic meta-analyses
[75,76], and these concepts can be extended to comparative analyses. The former are multi-
location experiments/observations designed to enable a subsequent synthesis, while the latter
is a comparative analysis that is continuously updated with new data. Such approaches
provide powerful ways to collect new data globally, expand the phylogenetic diversity of
taxa, and resolve major gaps in knowledge that are vital to address important eco-
evolutionary questions and inform conservation.
Concluding remarks

Comparative studies and meta-analyses are essential parts of modern research, revealing broad patterns in ecology and evolution. We assert that merging these methodologies into a unified framework will be transformative. Leveraging multilevel modelling and accounting for variation in sampling have the potential to shift understanding of biological variation. Testing hypotheses across different levels of organization will also illuminate variation within and between communities, and the importance of species interactions in driving trait variation. Moreover, the analysis of function-valued traits will broaden taxonomic coverage and may shape understanding of reaction norms (see Outstanding Questions). Assessing how the integration of comparative and meta-analyses will transform knowledge of macroevolutionary patterns will require large and complex data sets. Open science communities can expand current data collections (see Outstanding Questions) and undertake ambitious projects that will unlock the full potential of ecological and evolutionary syntheses.
Outstanding questions

- How does multilevel modelling change the relative contribution of evolutionary history in shaping macroevolutionary patterns?
- What is the magnitude of publication bias in comparative studies and meta-analyses, and how does this bias affect estimations of macroevolutionary patterns?
- How can sophisticated phylogenetic methods be integrated into meta-analytic models to enhance understanding of evolutionary history?
- How can reporting guidelines tailored specifically for phylogenetic comparative studies be developed to ensure robustness and transparency?
- To what extent do macroevolutionary patterns differ between species- and community-level analyses?
- To what extent can multivariate analysis of function-valued traits capture the shape of complex reaction norms?
- How effective are multivariate comparative models in estimating population and species-level traits when data are missing?
- How can open synthesis communities be harnessed to augment data compendia with information that is appropriate for use in multilevel models?
**BOX 1: Mathematical foundations of comparative and meta-analyses**

A typical phylogenetic comparative model can be formally described as:

\[ t_i = (Xb)_i + p_i + e_i \]

where \( t_i \) is a trait mean for species \( i \), \( Xb \) describes a series of ‘fixed’ effects impacting population level changes in trait means, \( p_i \) is the phylogenetic effect, assumed to be sampled from a normal distribution with a mean of 0 and covariance matrix proportional to the phylogenetic correlation matrix among taxa, \( C \), \( p \sim \mathcal{N}(0, \sigma^2_p C) \), and \( e_i \) is the residual effect \( e \sim \mathcal{N}(0, \sigma^2_e I) \). \( C \) is assumed to be known (estimated from phylogenetic tree) and variances are estimated. Such analyses ignore within-species variation and sampling error. Different models of evolution can be used to place restrictions on the \( C \) matrix to impact \( p_i \).

In contrast, a typical meta-analytic model in ecology and evolution can be described as:

\[ y_{ijk} = (Xb)_{ijk} + u_j + s_i + m_{ijk} + p_i + e_{ijk} \]

where \( y_{ijk} \) is the \( k \)-th standardized effect size from study \( j \) and species \( i \), \( u_j \) and \( s_i \) are the study- and non-phylogenetic species-specific effects, assumed to be sampled from multivariate normal distribution \( u \sim \mathcal{N}(0, \sigma^2_u I) \) and \( s \sim \mathcal{N}(0, \sigma^2_s I) \), respectively, and \( m_{ijk} \) is the known sampling error for the effect (calculated using sampling variance equations for effect sizes). While different models of evolution can be used on \( p_i \), meta-analyses typically do not investigate these patterns.

We can now explicitly merge typical comparative and meta-analytic models to provide the best of both worlds. We may still use trait means and covariates but include the sampling variance and within-species variation to decompose sources of variance. A phylogenetic multilevel meta-analysis might look instead as:
\[ t_{ijk} = (Xb)_{ijk} + u_j + s_i + m_{ijk} + p_i + e_{ijk} \]

Such a model allows us to understand how the trait evolves, provides opportunities to improve precision and decompose variance, thereby informing us about the possible explanatory factors that may be driving relationships.

We can then extend these concepts to the analysis of community-level patterns:

\[ b_{1,m} = (X_c b_c)_m + m_m + \epsilon_m \]

where \( b_1 \) is estimated for the \( m \)-th community from

\[ t_{ijk} = b_{0,m} + b_{1,m}x + u_j + s_i + m_{ijk} + p_i + e_{ijk}, \text{ and} \]

\[ \epsilon_m \sim N(0, \sigma^2_\epsilon D) \]

In the above two-level model \( b_{1,m} \) is the community-level parameter (e.g., slope or curve parameter) measured in community \( m \), \( b_{0,m} \) is the community-level intercept, \( D \) is the distance correlation matrix describing spatial autocorrelation between communities, \( e_{jkm} \) is the community-level residual sampled from \( N(0, \sigma^2_\epsilon I) \) and \( X_c b_c \) describes the between-community fixed effects. Sampling variance \( m_m \) is equal to the estimation error of the derived \( b_1 \) parameter. This example is simplified assuming estimation of only two parameters in each community (\( b_0 \) and \( b_1 \)), but similar logic can be applied to any coefficient of \( b \).

We can also extend these concepts to the analysis of function-valued traits, which take multiple parameters in a multivariate model. We define function-valued traits as traits that can be expressed as arbitrary functions (not necessarily linear) of one or multiple covariates.

In the simplest case, a comparative (meta-)analysis of a function-valued trait uses parameters (e.g., slope, curvature, optima) of the underlying function as responses. An example of a two-parameter model can be described as:
where $t_{ijk}^{(y)}$ and $t_{ijk}^{(z)}$ are parameters defining a function-valued trait, $G_p$ is the phylogenetic covariance matrix between traits $(y)$ and $(z)$, and $M^{(z)}$ is the matrix of sampling covariances for a given trait. This model can be extended to more than two parameters by following similar principles.

Examples to implement these approaches can be found at

https://szymekdr.github.io/meta_comparative_analysis/.
Glossary

- **Ancestral state/trait reconstruction**: The process of inferring the characteristics or traits of ancestors in a phylogenetic tree by analysing the distribution of traits in extant species and the patterns of trait evolution.

- **Effect measure**: Statistical metrics used to quantify the magnitude and direction of an effect or relationship observed in a study (e.g., association between two variables, comparison between two groups, trait mean), often used in comparative analyses and meta-analyses.

- **Effect size**: Standardized effect measure used in meta-analyses. Note that the term ‘effect size’ can also refer to the magnitude and direction of an observed effect or relationship between variables.

- **Meta-analysis**: Statistical method that combines effect sizes from multiple independent studies to obtain an overall estimate of an effect or relationship and its heterogeneity. Effect sizes are typically weighted based on a metric that reflects study quality (e.g., weighted by sample size or precision).

- **Multilevel framework**: An analytical approach that accounts for hierarchical structures in data, particularly when studying nested levels of organization, such as individuals within populations or species within communities.

- **Non-independence**: A situation where data points or observations are not statistically independent, which can lead to biased results if not properly accounted for in the analysis.

- **Phylogenetic comparative analysis**: Statistical method that incorporate the phylogenetic relationships among species to study evolutionary patterns and test hypotheses related to trait evolution and adaptation.
• **Phylogenetic generalized linear mixed models**: A statistical modelling approach that combines phylogenetic information with generalized linear mixed models to investigate the relationships between traits and other factors while accounting for phylogenetic non-independence and other sources of non-independence.

• **Phylogenetic multivariate meta-analysis**: Meta-analysis that incorporates multiple effect sizes simultaneously to quantify overall effects and effect sizes’ correlation, while also accounting for phylogenetic relatedness between species.

• **Phylogenetic non-independence**: Occurs when species are related through shared evolutionary history, potentially leading to correlations among data points that need to be addressed in comparative analyses.

• **Phylogenetic signal**: A measure indicating the degree to which the variation in traits among species reflects their phylogenetic relatedness.

• **Publication bias**: The tendency for published research to be biased towards statistically significant or positive results, leading to an overestimation of overall effects.

• **Reproducibility**: The ability to reproduce research findings using the same data, methods, and analyses, ensuring the reliability and validity of scientific results.

• **Sampling variance**: The variation in effect measures that result from variation in sampling effort, which is intricately linked to sample size.

• **Spatial dependence**: A condition where data points in space are not independent, leading to spatial autocorrelation that should be considered in analyses.

• **Systematic review**: A transparent, reproducible, objective, and rigorous review of the literature.
• **Taxonomic chauvinism**: An attitude or bias favoring certain taxonomic groups over others, potentially leading to overlooking important ecological or evolutionary information.

• **Time calibration**: The process of estimating the age of nodes to infer the timing of trait divergence and identify patterns of trait evolution.
Acknowledgements

This project is the result of a workshop organized with funding provided to PP from a UNSW Scientia PhD scholarship. We acknowledge the Bedegal, Dharawal, Darug, and Ngunnawal people, the traditional custodians of the land in which this work was conducted.

Author contributions

Conceptualization: All authors
Methodology: All authors
Software: SMD, DWAN, PP
Validation: SMD, DWAN, PP
Formal analysis: SMD, DWAN, PP
Investigation: SMD, DWAN, PP
Resources: None.
Data curation: SMD, DWAN
Writing - Original Draft: PP, FS, SN, DWAN, SB
Writing - Review & Editing: All authors
Visualization: NCW, ML, PP
Supervision: SN, SMD
Project administration: PP
Funding acquisition: None.

Declaration of interests

We declare not conflicts of interest.

Declaration of AI and AI-assisted technologies in the writing process

The authors declare having used GPT 4.0 and GPT 3.5 (OpenAI) to improve the readability and language of this manuscript. The authors have carefully reviewed and edited the content generated by these tools and take full responsibility for the content of the publication.
References


39. Yang, Y. *et al.* (2023) Publication bias impacts on effect size, statistical power, and magnitude (Type M) and sign (Type S) errors in ecology and evolutionary biology. *BMC Biology* 21, 71


64. Debastiani, V.J. et al. (2021) Using phylogenetic information to impute missing functional trait values in ecological databases. *Ecological Informatics* 63, 101315
Figure 1: Conceptual frameworks used to analyse comparative evolutionary and ecological variation.

A) Comparative analyses typically investigate evolutionary processes giving rise to trait differences at the tip of the phylogeny, yet these analyses are often limited to species-level (mean) values. B) Biological multilevel meta-analyses typically use highly heterogeneous datasets and partition the variance into different components to explain variation in effect sizes. These analyses also often
incorporate publication bias tests and follow rigorous reporting practices. C) The unified approach we propose merges the strength of both approaches. This approach improves quantifying and decomposing ecological, methodological, and evolutionary variation in biological datasets.

Figure 2: Conceptual frameworks for community-level and function-valued analyses. A: Multilevel comparative analyses can be used to investigate questions at broader scales, by combining results from spatially dependent models performed at the community level. B: Function-valued analyses use multivariate analyses to investigate patterns from multiple parameters of a continuous trait (e.g., thermal performance curve, TPC). This approach can leverage datasets with missing data, and investigate overall effects for each parameter, as well as trait covariation and trade-offs. CTmin: critical thermal maximum; Topt: thermal optimum; CTmax: critical thermal maximum.