

1 **New horizons for comparative studies and meta-analyses**

2 Patrice Pottier^{1*}, Daniel W.A. Noble²⁺, Frank Seebacher³⁺, Nicholas C. Wu⁴⁺, Samantha Burke¹,

3 Malgorzata Lagisz^{1,5}, Lisa E. Schwanz¹, Szymon M. Drobniak^{1,6#}, Shinichi Nakagawa^{1,5#}

4

5 ¹ Evolution & Ecology Centre, School of Biological, Earth and Environmental Sciences, UNSW,
6 Sydney, Australia.

7 ²Division of Ecology and Evolution, Research School of Biology, College of Science, The Australian
8 National University, Canberra, Australian Capital Territory, Australia.

9 ³School of Life and Environmental Sciences, University of Sydney, Sydney, New South Wales 2006,
10 Australia.

11 ⁴Hawkesbury Institute for the Environment, Western Sydney University, NSW 2753, Australia

12 ⁵Theoretical Sciences Visiting Program, Okinawa Institute of Science and Technology Graduate
13 University, Onna, 904-0495, Japan.

14 ⁶Institute of Environmental Sciences, Jagiellonian University, Krakow, Poland.

15 *Corresponding author: Pottier, P. (p.pottier@unsw.edu.au); Twitter: @PatriceEcoEvo

16 #These authors supervised this work equally

17 +These authors contributed equally

18 **ORCID**

19 Patrice Pottier <https://orcid.org/0000-0003-2106-6597>

20 Daniel W.A. Noble <https://orcid.org/0000-0001-9460-8743>

21 Frank Seebacher <https://orcid.org/0000-0002-2281-9311>

22 Nicholas C. Wu <https://orcid.org/0000-0002-7130-1279>

23 Samantha Burke <https://orcid.org/0000-0001-6902-974X>

24 Malgorzata Lagisz <https://orcid.org/0000-0002-3993-6127>

25 Lisa E. Schwanz <https://orcid.org/0000-0001-5864-7112>

26 Szymon M. Drobniak <https://orcid.org/0000-0001-8101-6247>

27 Shinichi Nakagawa <https://orcid.org/0000-0002-7765-5182>

28 **Highlights**

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

31

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

34

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

38

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

42

43

44

45

46

47

48

49

50

51

52 **Abstract**

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

64

65 **Keywords:** multilevel modelling, multivariate analysis, phylogenetic generalized linear
66 mixed model, sampling variance, phylogenetic signal

67

68

69

70

71

72

73 **History and purpose of comparative studies and meta-analyses**

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

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

93 Similar to comparative analyses, biological meta-analyses (see Glossary) typically
94 compare traits across taxa. Phylogenetic correction is therefore essential here too. However,
95 unlike comparative analyses, meta-analyses typically do not ask explicit questions about
96 evolutionary processes and are primarily focussed on present-day phenotypes [15]. Meta-

97 analysis was developed in medicine and social sciences in parallel with evolutionary
98 comparative analyses [16,17]. The purpose of early meta-analyses was to synthesize results
99 across similar studies to detect the overall magnitude of treatment effects in a single species
100 (humans). Eventually, meta-analyses were extended to ecology and evolution to integrate
101 data from multiple species [18]. Their purpose shifted from simple quantification of effects to
102 establishing the state-of-knowledge in a field and reassessing established hypotheses with
103 mixed empirical support (e.g., [19,20]). Contrary to traditional comparative analyses,
104 biological meta-analyses purposefully use heterogeneous datasets to estimate effect variation
105 among multiple sources (e.g., different populations, species, geographical areas), while
106 explicitly accounting for the variation due to sampling effort [17]. Meta-analyses also provide
107 ways to test for publication biases. Such analyses not only assess the reliability of research
108 findings but can also illuminate social dimensions in the research and publishing process
109 [21].

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

122 **A unified approach for comparative studies**

123 **A multilevel framework to decompose biological and methodological variation**

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

147 Ornstein-Uhlenbek). Implementing these methods can help decipher the evolutionary
148 processes that led to present-day phenotypes - sparking new hypotheses and pioneering novel
149 research avenues.

150

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

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

171

172 **Integrating different effect size measures for greater flexibility**

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

189

190 **Reporting guidelines to promote transparency and reproducibility**

191 Another key distinction between comparative studies and meta-analyses pertains to
192 the methods and reporting used to ensure reproducibility. The historical connectivity of meta-
193 analysis with medicine and social sciences has generated guidelines and recommendations to
194 ensure transparency and reproducibility [48,49]. Adopting reporting standards aids in study
195 design, ensures the inclusion of important methodological details, and ultimately elevates the
196 reliability of research. In ecology and evolutionary biology, established guidelines for

197 systematic reviews and meta-analyses, such as PRISMA-EcoEvo [50], can be adopted with
198 little adjustments. However, there is a need to conceptualize reporting guidelines tailored
199 specifically for phylogenetic comparative studies. Comparative studies may also benefit from
200 using best practices developed for meta-analyses, as both approaches share similar tools and
201 objectives, as we have argued so far. Ideally, literature-based data syntheses should stem
202 from a systematic review (see Glossary) with fully documented and reproducible procedures
203 [51]. This practice simplifies updates and replications which helps build more trust in
204 evidence [52]. However, this is not always feasible, particularly because most data used in
205 comparative studies are taken from data compendia [53]. Building upon previous data
206 compilations to incorporate important information (i.e., data provenance, metadata, sampling
207 variance) could elevate the robustness of future studies and broaden the applicability of data
208 compendia for comparative- and meta-analyses.

209

210 **New opportunities and directions for comparative studies**

211 **Community-level responses and eco-evolutionary patterns**

212 Community- and ecosystem-level analyses are routinely used in plant and community
213 ecology. Yet, use of community-level responses is limited in phylogenetic comparative
214 analyses. This is perhaps because phylogenetic components cannot be modelled at the
215 community level, though phylogeny can still be modelled within each community. For
216 instance, Markovski and colleagues [54] recently investigated global variation in the
217 relationship between population size and sexual dimorphism. The authors quantified
218 community-level standardized slopes where a positive relationship between sexual
219 dimorphism and population size is predicted if sexual selection promotes viability via good
220 genes [55]. The study yielded 2,592 slopes and error variances from phylogenetically
221 controlled analyses in each community (grid cell), and then estimated overall slopes for both

222 resident and migratory species. Notably, this study controlled for spatial dependence (see
223 Glossary) across grid cells and decomposed sources of variance. The researchers found the
224 expected benefits of sexual selection in resident species, but not in migratory. This is an
225 elegant example of how marrying phylogenetic comparative analysis with spatially-
226 controlled meta-analysis can lead to significant new insights. At the community level, species
227 assemblages vary greatly, which provides interesting insights into ecological (e.g. spatial
228 variation) and evolutionary (e.g., phylogenetic signal) processes governing biological
229 variation at different scales. In addition, studying community-level patterns allow to capture
230 the influence of species interactions that may be missed in individual- or species-level
231 analyses. Such interactions (e.g., predator-prey, competition) are integral parts of an
232 ecosystem and profoundly influence species traits within each community. Community-level
233 analyses can also illuminate variability in species' responses to environmental change across
234 communities. This variability can be key in identifying species and communities more
235 sensitive to disturbance.

236

237 **Function-valued traits and multivariate meta-analyses**

238 Function-valued traits are organismal responses to continuous variables such as
239 temperature, pH, or age [56,57]. Examples include performance curves, growth trajectories,
240 reflectance spectra, or sonograms. Function-valued traits can perhaps be seen as a new type
241 of effect measure that combines a set of parameters. Complex function-valued traits are better
242 depicted by curves than lines, and such traits can be summarized in various descriptors
243 (parameters) of the curve (e.g., intercept, slope, peak, asymptote). The best way to model
244 multi-parameter measures (i.e., function-valued traits) is with phylogenetic multivariate meta-
245 analysis (see Glossary) (PMMA; [23]), which can estimate not only correlations between
246 parameters of function-valued traits, but also accounts for uncertainty in measurements.

247 Notably, PMMA is an extension of phylogenetic comparative multivariate analysis, which is
248 used to examine, for example, morphometric data where ‘landmarks’ describe complex traits
249 such as body shapes [58–60].

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

263 Another major benefit of using multivariate models for the analysis of function-
264 valued traits is their potential to improve precision. When multiple parameters of the
265 function-valued traits are correlated, precision around parameters can be improved by
266 explicitly accounting for the covariance among variables (“borrowing of strength” [63]).
267 Using multivariate models also means that some of the parameters of function-valued traits
268 can be missing as long as not all trait values are missing for each species [64,65]. Indeed,
269 phylogenetic multivariate models and related techniques can impute missing data, as missing
270 trait values are inferred from the available parameters. Data imputation will extend not only
271 the number of traits but also the number of species that can be examined, although the

272 effective use of imputation remains uncommon [62,66]. Taken together, the analysis of
273 function-valued traits could dramatically increase the number of biological questions that can
274 be asked, and better capture the intricate shape of biological responses.

275

276 **Towards next-generation comparative studies via open synthesis communities**

277 Community-level and function-valued analyses may require much larger datasets than
278 ordinary analyses. This calls for scientific “community-level” collaboration [67]. Movements
279 towards big-team science are already happening in the form of global research networks and
280 globally distributed experiments such as SPI-Birds [68], the Global Urban Evolution Project
281 [69], the Nutrient Network [70], and large-scale collaborative databases such as BioTIME
282 [71] or PREDICTS [72].

283 A rapid spread of global research networks provides a foundation for “open synthesis
284 communities”, where scientists with similar interests can plan and conduct comparative and
285 meta-analyses together with research synthesis specialists (e.g., librarians and information
286 scientists [73]). Simultaneously, such communities increasingly adhere to the principles of
287 open science, embracing open participation, materials, data, and code [74]. An open synthesis
288 community can carry out prospective meta-analyses as well as living/dynamic meta-analyses
289 [75,76], and these concepts can be extended to comparative analyses. The former are multi-
290 location experiments/observations designed to enable a subsequent synthesis, while the latter
291 is a comparative analysis that is continuously updated with new data. Such approaches
292 provide powerful ways to collect new data globally, expand the phylogenetic diversity of
293 taxa, and resolve major gaps in knowledge that are vital to address important eco-
294 evolutionary questions and inform conservation.

295

296

297 **Concluding remarks**

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

311

312

313

314

315

316

317

318

319

320

321

322 **Outstanding questions**

- 323 • How does multilevel modelling change the relative contribution of evolutionary
324 history in shaping macroevolutionary patterns?
- 325 • What is the magnitude of publication bias in comparative studies and meta-analyses,
326 and how does this bias affect estimations of macroevolutionary patterns?
- 327 • How can sophisticated phylogenetic methods be integrated into meta-analytic models
328 to enhance understanding of evolutionary history?
- 329 • How can reporting guidelines tailored specifically for phylogenetic comparative
330 studies be developed to ensure robustness and transparency?
- 331 • To what extent do macroevolutionary patterns differ between species- and
332 community-level analyses?
- 333 • To what extent can multivariate analysis of function-valued traits capture the shape of
334 complex reaction norms?
- 335 • How effective are multivariate comparative models in estimating population and
336 species-level traits when data are missing?
- 337 • How can open synthesis communities be harnessed to augment data compendia with
338 information that is appropriate for use in multilevel models?

339

340

341

342

343

344

345

346

347 **BOX 1: Mathematical foundations of comparative and meta-analyses**

348 A typical phylogenetic comparative model can be formally described as:

349
$$t_i = (\mathbf{Xb})_i + p_i + e_i$$

350 where t_i is a trait mean for species i , \mathbf{Xb} describes a series of ‘fixed’ effects impacting
351 population level changes in trait means, p_i is the phylogenetic effect, assumed to be sampled
352 from a normal distribution with a mean of 0 and covariance matrix proportional to the
353 phylogenetic correlation matrix among taxa, \mathbf{C} , $\mathbf{p} \sim \mathcal{N}(\mathbf{0}, \sigma_p^2 \mathbf{C})$, and e_i is the residual effect
354 $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \sigma_e^2 \mathbf{I})$. \mathbf{C} is assumed to be known (estimated from phylogenetic tree) and variances
355 are estimated. Such analyses ignore within-species variation and sampling error. Different
356 models of evolution can be used to place restrictions on the \mathbf{C} matrix to impact p_i .

357

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

359
$$y_{ijk} = (\mathbf{Xb})_{ijk} + u_j + s_i + m_{ijk} + p_i + e_{ijk}$$

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

366

367 We can now explicitly merge typical comparative and meta-analytic models to provide the
368 best of both worlds. We may still use trait means and covariates but include the sampling
369 variance and within-species variation to decompose sources of variance. A phylogenetic
370 multilevel meta-analysis might look instead as:

371
$$t_{ijk} = (\mathbf{X}\mathbf{b})_{ijk} + u_j + s_i + m_{ijk} + p_i + e_{ijk}$$

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

375

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

377
$$b_{1,m} = (\mathbf{X}_c \mathbf{b}_c)_m + m_m + \varepsilon_m$$

378 where b_1 is estimated for the m -th community from

379
$$t_{ijk} = b_{0,m} + b_{1,m}x + u_j + s_i + m_{ijk} + p_i + e_{ijk}, \text{ and}$$

380
$$\varepsilon_m \sim N(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{D})$$

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

388

389 We can also extend these concepts to the analysis of function-valued traits, which take
 390 multiple parameters in a multivariate model. We define function-valued traits as traits that
 391 can be expressed as arbitrary functions (not necessarily linear) of one or multiple covariates.
 392 In the simplest case, a comparative (meta-)analysis of a function-valued trait uses parameters
 393 (e.g., slope, curvature, optima) of the underlying function as responses. An example of a two-
 394 parameter model can be described as:

395
$$\begin{pmatrix} t_{ijk}^{(y)} \\ t_{ijk}^{(z)} \end{pmatrix} = \begin{pmatrix} (\mathbf{Xb}^{(y)})_{ijk} + u_j^{(y)} + s_i^{(y)} + m_{ijk}^{(y)} + p_i^{(y)} + e_{ijk}^{(y)} \\ (\mathbf{Xb}^{(z)})_{ijk} + u_j^{(z)} + s_i^{(z)} + m_{ijk}^{(z)} + p_i^{(z)} + e_{ijk}^{(z)} \end{pmatrix}$$

396
$$(\mathbf{p}^{(y)'}, \mathbf{p}^{(z)'}) \sim \mathcal{N}(\mathbf{0}, \mathbf{G}_p \otimes \mathbf{C})$$

397
$$(\mathbf{m}^{(y)'}, \mathbf{m}^{(z)'}) \sim \mathcal{N}(\mathbf{0}, \mathbf{M}^{(y)} \oplus \mathbf{M}^{(z)})$$

398 where $t_{ijk}^{(y)}$ and $t_{ijk}^{(z)}$ are parameters defining a function-valued trait, \mathbf{G}_p is the phylogenetic
 399 covariance matrix between traits (y) and (z), and $\mathbf{M}^{(\cdot)}$ is the matrix of sampling covariances
 400 for a given trait. This model can be extended to more than two parameters by following
 401 similar principles.

402 Examples to implement these approaches can be found at

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

404

405

406

407

408

409

410

411

412

413

414

415

416

417 **Glossary**

- 418 • *Ancestral state/trait reconstruction*: The process of inferring the characteristics or
419 traits of ancestors in a phylogenetic tree by analysing the distribution of traits in
420 extant species and the patterns of trait evolution.
- 421 • *Effect measure*: Statistical metrics used to quantify the magnitude and direction of an
422 effect or relationship observed in a study (e.g., association between two variables,
423 comparison between two groups, trait mean), often used in comparative analyses and
424 meta-analyses.
- 425 • *Effect size*: Standardized effect measure used in meta-analyses. Note that the term
426 ‘effect size’ can also refer to the magnitude and direction of an observed effect or
427 relationship between variables.
- 428 • *Meta-analysis*: Statistical method that combines effect sizes from multiple
429 independent studies to obtain an overall estimate of an effect or relationship and its
430 heterogeneity. Effect sizes are typically weighted based on a metric that reflects study
431 quality (e.g., weighted by sample size or precision).
- 432 • *Multilevel framework*: An analytical approach that accounts for hierarchical structures
433 in data, particularly when studying nested levels of organization, such as individuals
434 within populations or species within communities.
- 435 • *Non-independence*: A situation where data points or observations are not statistically
436 independent, which can lead to biased results if not properly accounted for in the
437 analysis.
- 438 • *Phylogenetic comparative analysis*: Statistical method that incorporate the
439 phylogenetic relationships among species to study evolutionary patterns and test
440 hypotheses related to trait evolution and adaptation.

- 441 • *Phylogenetic generalized linear mixed models*: A statistical modelling approach that
442 combines phylogenetic information with generalized linear mixed models to
443 investigate the relationships between traits and other factors while accounting for
444 phylogenetic non-independence and other sources of non-independence.
- 445 • *Phylogenetic multivariate meta-analysis*: Meta-analysis that incorporates multiple
446 effect sizes simultaneously to quantify overall effects and effect sizes' correlation,
447 while also accounting for phylogenetic relatedness between species.
- 448 • *Phylogenetic non-independence*: Occurs when species are related through shared
449 evolutionary history, potentially leading to correlations among data points that need to
450 be addressed in comparative analyses.
- 451 • *Phylogenetic signal*: A measure indicating the degree to which the variation in traits
452 among species reflects their phylogenetic relatedness.
- 453 • *Publication bias*: The tendency for published research to be biased towards
454 statistically significant or positive results, leading to an overestimation of overall
455 effects.
- 456 • *Reproducibility*: The ability to reproduce research findings using the same data,
457 methods, and analyses, ensuring the reliability and validity of scientific results.
- 458 • *Sampling variance*: The variation in effect measures that result from variation in
459 sampling effort, which is intricately linked to sample size.
- 460 • *Spatial dependence*: A condition where data points in space are not independent,
461 leading to spatial autocorrelation that should be considered in analyses.
- 462 • *Systematic review*: A transparent, reproducible, objective, and rigorous review of the
463 literature.

- 464 • *Taxonomic chauvinism*: An attitude or bias favoring certain taxonomic groups over
465 others, potentially leading to overlooking important ecological or evolutionary
466 information.
- 467 • *Time calibration*: The process of estimating the age of nodes to infer the timing of
468 trait divergence and identify patterns of trait evolution.

469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488

489 **Acknowledgements**

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

493 **Author contributions**

494 Conceptualization: All authors

495 Methodology: All authors

496 Software: SMD, DWAN, PP

497 Validation: SMD, DWAN, PP

498 Formal analysis: SMD, DWAN, PP

499 Investigation: SMD, DWAN, PP

500 Resources: None.

501 Data curation: SMD, DWAN

502 Writing - Original Draft: PP, FS, SN, DWAN, SB

503 Writing - Review & Editing: All authors

504 Visualization: NCW, ML, PP

505 Supervision: SN, SMD

506 Project administration: PP

507 Funding acquisition: None.

508 **Declaration of interests**

509 We declare not conflicts of interest.

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

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

514 **References**

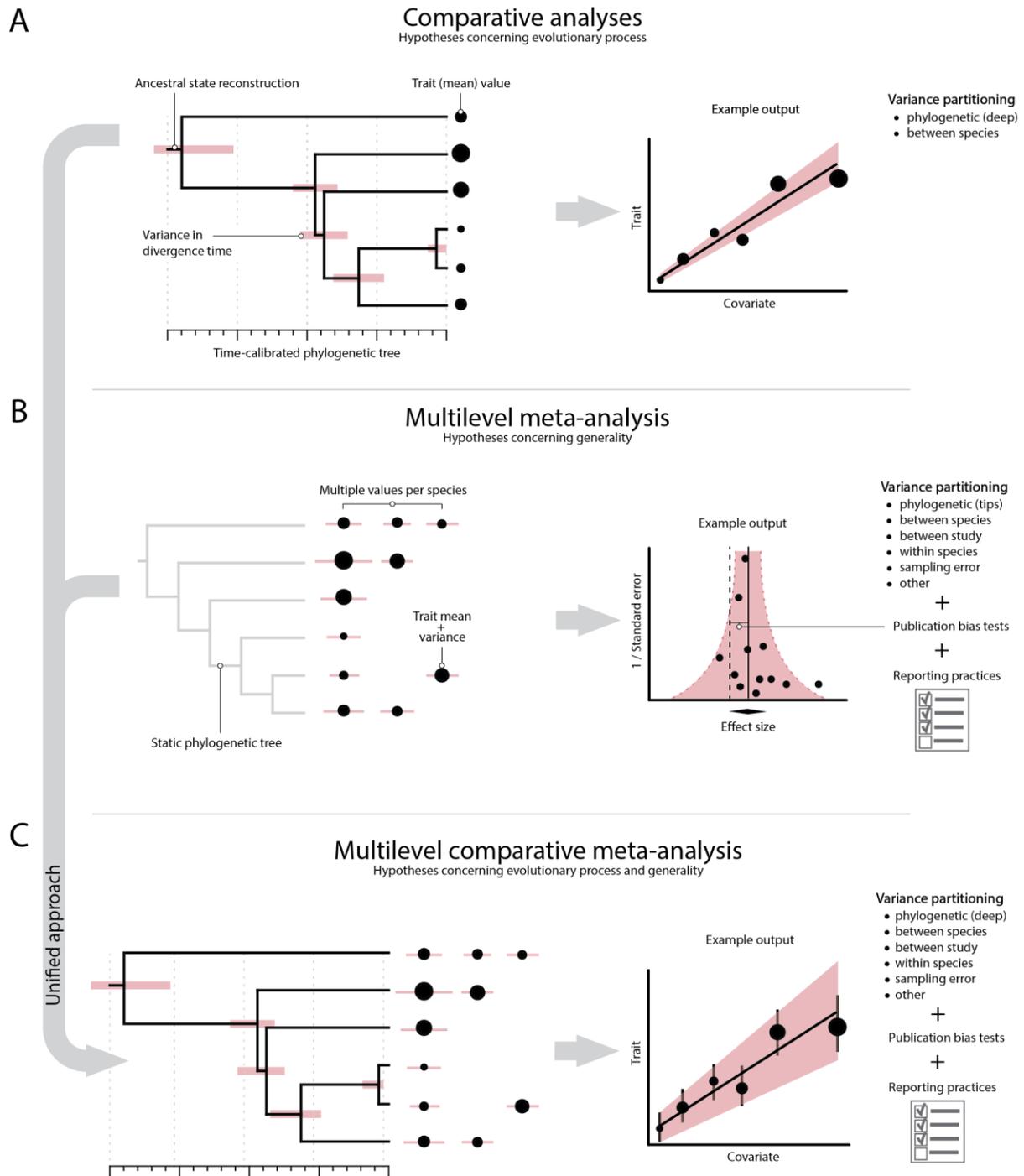
- 515 1. Benton, M.J. (2016) Origins of Biodiversity. *PLOS Biology* 14, e2000724
- 516 2. Millar, J. (2021) Roman Climate Awareness in Pliny the Elder’s Natural History.
517 *Classical Antiquity* 40, 249–282
- 518 3. Harvey, P.H. and Purvis, A. (1991) Comparative methods for explaining adaptations.
519 *Nature* 351, 619–624
- 520 4. He, J. *et al.* (2023) A global assessment of Bergmann’s rule in mammals and birds.
521 *Global Change Biology* 29, 5199–5210
- 522 5. Schmidt-Nielsen, K. (1975) Scaling in biology: The consequences of size. *Journal of*
523 *Experimental Zoology* 194, 287–307
- 524 6. Felsenstein, J. (1985) Phylogenies and the Comparative Method. *The American*
525 *Naturalist* 125, 1–15
- 526 7. Pagel, M. (1999) Inferring the historical patterns of biological evolution. *Nature* 401,
527 877–884
- 528 8. Blomberg, S.P. *et al.* (2003) Testing for Phylogenetic Signal in Comparative Data:
529 Behavioral Traits Are More Labile. *Evolution* 57, 717–745
- 530 9. Garland, T., Jr *et al.* (2005) Phylogenetic approaches in comparative physiology. *Journal*
531 *of Experimental Biology* 208, 3015–3035
- 532 10. Weber, M.G. and Agrawal, A.A. (2012) Phylogeny, ecology, and the coupling of
533 comparative and experimental approaches. *Trends in Ecology & Evolution* 27, 394–403
- 534 11. Garamszegi, L.Z. (2014) *Modern Phylogenetic Comparative Methods and Their*
535 *Application in Evolutionary Biology: Concepts and Practice*, Springer
- 536 12. Liedtke, H.C. *et al.* (2022) The evolution of reproductive modes and life cycles in
537 amphibians. *Nature Communications* 13, 7039
- 538 13. Ho, S.Y.W. *et al.* (2011) Time-dependent rates of molecular evolution. *Molecular*
539 *Ecology* 20, 3087–3101
- 540 14. Alfaro, M.E. *et al.* (2009) Nine exceptional radiations plus high turnover explain species
541 diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences* 106,
542 13410–13414
- 543 15. Chamberlain, S.A. *et al.* (2012) Does phylogeny matter? Assessing the impact of
544 phylogenetic information in ecological meta-analysis. *Ecology Letters* 15, 627–636
- 545 16. Arnqvist, G. and Wooster, D. (1995) Meta-analysis: synthesizing research findings in
546 ecology and evolution. *Trends in Ecology & Evolution* 10, 236–240

- 547 17. Gurevitch, J. *et al.* (2018) Meta-analysis and the science of research synthesis. *Nature*
548 555, 175–182
- 549 18. Gurevitch, J. *et al.* (1992) A Meta-Analysis of Competition in Field Experiments. *The*
550 *American Naturalist* 140, 539–572
- 551 19. Prokop, Z.M. *et al.* (2012) Meta-analysis suggests choosy females get sexy sons more
552 than “good genes.” *Evolution* 66, 2665–2673
- 553 20. Lefevre, S. (2016) Are global warming and ocean acidification conspiring against marine
554 ectotherms? A meta-analysis of the respiratory effects of elevated temperature, high CO₂
555 and their interaction. *Conservation Physiology* 4
- 556 21. Marks-Anglin, A. and Chen, Y. (2020) A historical review of publication bias. *Research*
557 *Synthesis Methods* 11, 725–742
- 558 22. Felsenstein, J. (2008) Comparative Methods with Sampling Error and Within-Species
559 Variation: Contrasts Revisited and Revised. *The American Naturalist* 171, 713–725
- 560 23. Nakagawa, S. and Santos, E.S.A. (2012) Methodological issues and advances in
561 biological meta-analysis. *Evolutionary Ecology* 26, 1253–1274
- 562 24. Hadfield, J.D. and Nakagawa, S. (2010) General quantitative genetic methods for
563 comparative biology: phylogenies, taxonomies and multi-trait models for continuous and
564 categorical characters. *Journal of Evolutionary Biology* 23, 494–508
- 565 25. Ives, A.R. and Helmus, M.R. (2011) Generalized linear mixed models for phylogenetic
566 analyses of community structure. *Ecological Monographs* 81, 511–525
- 567 26. Housworth, E.A. *et al.* (2004) The Phylogenetic Mixed Model. *The American Naturalist*
568 163, 84–96
- 569 27. Gallinat, A.S. and Pearse, W.D. (2021) Phylogenetic generalized linear mixed modeling
570 presents novel opportunities for eco-evolutionary synthesis. *Oikos* 130, 669–679
- 571 28. Ives, A.R. *et al.* (2007) Within-Species Variation and Measurement Error in Phylogenetic
572 Comparative Methods. *Systematic Biology* 56, 252–270
- 573 29. Garamszegi, L.Z. (2014) Uncertainties Due to Within-Species Variation in Comparative
574 Studies: Measurement Errors and Statistical Weights. In *Modern Phylogenetic*
575 *Comparative Methods and Their Application in Evolutionary Biology: Concepts and*
576 *Practice* (Garamszegi, L. Z., ed), pp. 157–199, Springer
- 577 30. Garamszegi, L.Z. and Møller, A.P. (2010) Effects of sample size and intraspecific
578 variation in phylogenetic comparative studies: a meta-analytic review. *Biological*
579 *Reviews* 85, 797–805

- 580 31. Garamszegi, L.Z. and Møller, A.P. (2017) Partitioning within-species variance in
581 behaviour to within- and between-population components for understanding evolution.
582 *Ecology Letters* 20, 599–608
- 583 32. Cinar, O. *et al.* Phylogenetic multilevel meta-analysis: A simulation study on the
584 importance of modelling the phylogeny. *Methods in Ecology and Evolution* 13, 383–395
- 585 33. Stone, G.N. *et al.* (2011) Controlling for non-independence in comparative analysis of
586 patterns across populations within species. *Philosophical Transactions of the Royal
587 Society B: Biological Sciences* 366, 1410–1424
- 588 34. Noble, D.W.A. *et al.* (2017) Nonindependence and sensitivity analyses in ecological and
589 evolutionary meta-analyses. *Molecular Ecology* 26, 2410–2425
- 590 35. Morrissey, M.B. (2016) Meta-analysis of magnitudes, differences and variation in
591 evolutionary parameters. *Journal of Evolutionary Biology* 29, 1882–1904
- 592 36. Bonnet, X. *et al.* (2002) Taxonomic chauvinism. *Trends in Ecology & Evolution* 17, 1–3
- 593 37. Garamszegi, L.Z. and Møller, A.P. (2012) Untested assumptions about within-species
594 sample size and missing data in interspecific studies. *Behavioral Ecology and
595 Sociobiology* 66, 1363–1373
- 596 38. Garamszegi, L.Z. and Møller, A.P. (2011) Nonrandom Variation in Within-Species
597 Sample Size and Missing Data in Phylogenetic Comparative Studies. *Systematic Biology*
598 60, 876–880
- 599 39. Yang, Y. *et al.* (2023) Publication bias impacts on effect size, statistical power, and
600 magnitude (Type M) and sign (Type S) errors in ecology and evolutionary biology. *BMC
601 Biology* 21, 71
- 602 40. Rosenthal, R. (1979) The file drawer problem and tolerance for null results.
603 *Psychological Bulletin* 86, 638–641
- 604 41. Nakagawa, S. *et al.* (2022) Methods for testing publication bias in ecological and
605 evolutionary meta-analyses. *Methods in Ecology and Evolution* 13, 4–21
- 606 42. Hedges, L.V. *et al.* (2012) A standardized mean difference effect size for single case
607 designs. *Research Synthesis Methods* 3, 224–239
- 608 43. Hedges, L.V. *et al.* (1999) The Meta-Analysis of Response Ratios in Experimental
609 Ecology. *Ecology* 80, 1150–1156
- 610 44. Nakagawa, S. *et al.* (2015) Meta-analysis of variation: ecological and evolutionary
611 applications and beyond. *Methods in Ecology and Evolution* 6, 143–152
- 612 45. Bland, J.M. and Altman, D.G. (2000) The odds ratio. *BMJ* 320, 1468
- 613 46. Hedges, L.V. and Olkin, I. (1985) *Statistical Methods for Meta-Analysis*, Academic Press

- 614 47. Nakagawa, S. and Cuthill, I.C. (2007) Effect size, confidence interval and statistical
615 significance: a practical guide for biologists. *Biological Reviews* 82, 591–605
- 616 48. Moher, D. *et al.* (2009) Preferred Reporting Items for Systematic Reviews and Meta-
617 Analyses: The PRISMA Statement. *Annals of Internal Medicine* 151, 264–269
- 618 49. Haddaway, N.R. *et al.* (2018) ROSES RepOrting standards for Systematic Evidence
619 Syntheses: pro forma, flow-diagram and descriptive summary of the plan and conduct of
620 environmental systematic reviews and systematic maps. *Environmental Evidence* 7, 7
- 621 50. O’Dea, R.E. *et al.* (2021) Preferred reporting items for systematic reviews and meta-
622 analyses in ecology and evolutionary biology: a PRISMA extension. *Biological Reviews*
623 96, 1695–1722
- 624 51. Foo, Y.Z. *et al.* (2021) A practical guide to question formation, systematic searching and
625 study screening for literature reviews in ecology and evolution. *Methods in Ecology and*
626 *Evolution* 12, 1705–1720
- 627 52. Parker, T.H. *et al.* (2016) Transparency in Ecology and Evolution: Real Problems, Real
628 Solutions. *Trends in Ecology & Evolution* 31, 711–719
- 629 53. Schwanz, L.E. *et al.* (2022) Best practices for building and curating databases for
630 comparative analyses. *Journal of Experimental Biology* 225, jeb243295
- 631 54. Markovski, *et al.* (2023) A global analysis reveals the dynamic relationship between
632 sexual selection and population abundance in space and time. *EcoEvoRxiv*.
633 <https://doi.org/10.32942/X2Q606>
- 634 55. Zahavi, A. (1975) Mate selection—A selection for a handicap. *Journal of Theoretical*
635 *Biology* 53, 205–214
- 636 56. Gomulkiewicz, R. *et al.* (2018) Variation and Evolution of Function-Valued Traits.
637 *Annual Review of Ecology, Evolution, and Systematics* 49, 139–164
- 638 57. Stinchcombe, J.R. and Kirkpatrick, M. (2012) Genetics and evolution of function-valued
639 traits: understanding environmentally responsive phenotypes. *Trends in Ecology &*
640 *Evolution* 27, 637–647
- 641 58. Adams, D.C. and Collyer, M.L. (2019) Phylogenetic Comparative Methods and the
642 Evolution of Multivariate Phenotypes. *Annual Review of Ecology, Evolution, and*
643 *Systematics* 50, 405–425
- 644 59. Adams, D.C. and Collyer, M.L. (2018) Multivariate Phylogenetic Comparative Methods:
645 Evaluations, Comparisons, and Recommendations. *Systematic Biology* 67, 14–31
- 646 60. Halliwell, B. *et al.* (2023) Multi-Response Phylogenetic Mixed Models: Concepts and
647 Application. *BioRxiv*, <https://doi.org/10.1101/2022.12.13.520338>

- 648 61. Westoby, M. *et al.* (2023) Phylogenetically conservative trait correlation: Quantification
649 and interpretation. *Journal of Ecology*
- 650 62. Pettersen, A. *et al.* (2023) Maternal behavioral thermoregulation facilitated evolutionary
651 transitions from egg laying to live birth. *Evolution Letters*, qrad031
- 652 63. Jackson, D. *et al.* (2017) Borrowing of strength and study weights in multivariate and
653 network meta-analysis. *Statistical Methods in Medical Research* 26, 2853–2868
- 654 64. Debastiani, V.J. *et al.* (2021) Using phylogenetic information to impute missing
655 functional trait values in ecological databases. *Ecological Informatics* 63, 101315
- 656 65. Noble, D.W.A. and Nakagawa, S. (2021) Planned missing data designs and methods:
657 Options for strengthening inference, increasing research efficiency and improving animal
658 welfare in ecological and evolutionary research. *Evolutionary Applications* 14, 1958–
659 1968
- 660 66. Callaghan, C.T. *et al.* (2021) Global abundance estimates for 9,700 bird species.
661 *Proceedings of the National Academy of Sciences* 118, e2023170118
- 662 67. Coles, N.A. *et al.* (2022) Build up big-team science. *Nature* 601, 505–507
- 663 68. Culina, A. *et al.* (2021) Connecting the data landscape of long-term ecological studies:
664 The SPI-Birds data hub. *Journal of Animal Ecology* 90, 2147–2160
- 665 69. Santangelo, J.S. *et al.* (2022) Global urban environmental change drives adaptation in
666 white clover. *Science* 375, 1275–1281
- 667 70. Borer, E.T. *et al.* (2014) Finding generality in ecology: a model for globally distributed
668 experiments. *Methods in Ecology and Evolution* 5, 65–73
- 669 71. Dornelas, M. *et al.* (2018) BioTIME: A database of biodiversity time series for the
670 Anthropocene. *Global Ecology and Biogeography* 27, 760–786
- 671 72. Hudson, L.N. *et al.* (2014) The PREDICTS database: a global database of how local
672 terrestrial biodiversity responds to human impacts. *Ecology and Evolution* 4, 4701–4735
- 673 73. Nakagawa, S. *et al.* (2020) A new ecosystem for evidence synthesis. *Nature Ecology &*
674 *Evolution* 4, 498–501
- 675 74. Haddaway, N.R. (2018) Open Synthesis: on the need for evidence synthesis to embrace
676 Open Science. *Environmental Evidence* 7, 26
- 677 75. Shackelford, G.E. *et al.* (2021) Dynamic meta-analysis: a method of using global
678 evidence for local decision making. *BMC Biology* 19, 33
- 679 76. Seidler, A.L. *et al.* (2019) A guide to prospective meta-analysis. *BMJ* 367, 15342
680



681

682

683 **Figure 1:** Conceptual frameworks used to analyse comparative evolutionary and ecological variation.

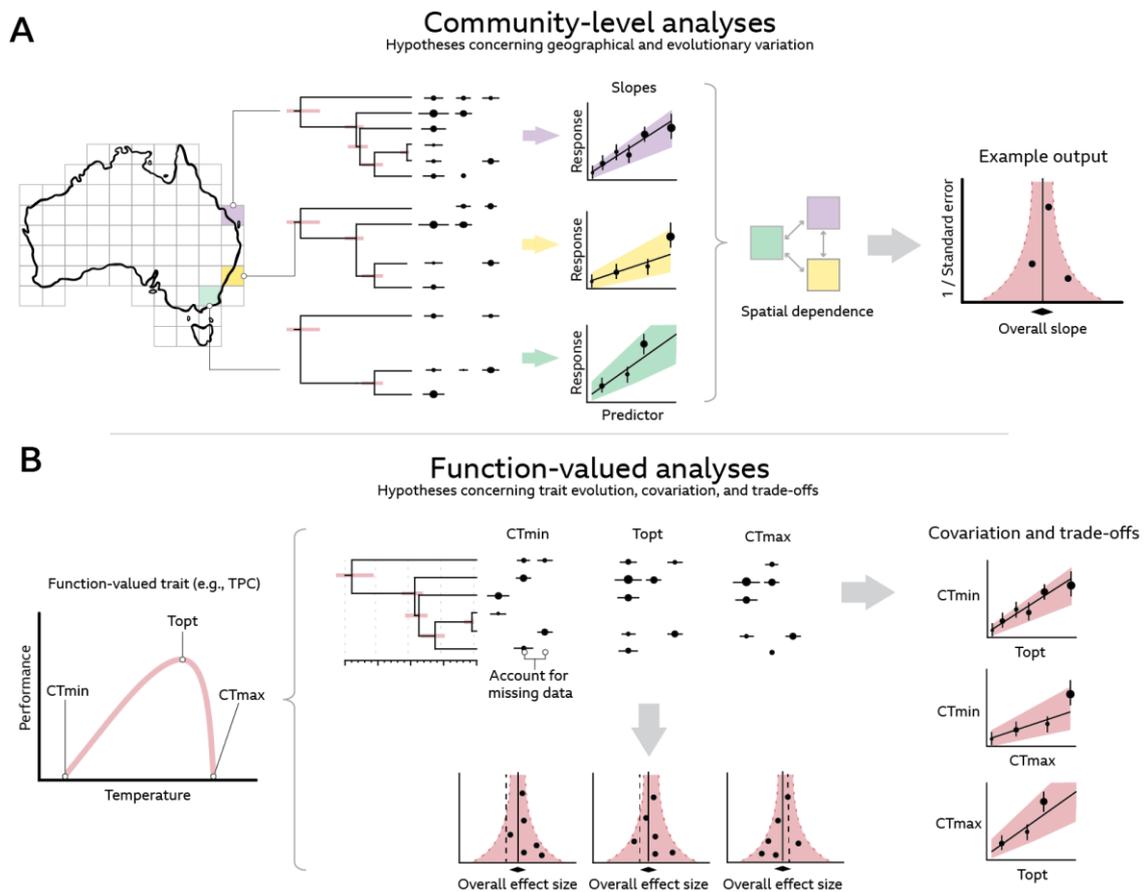
684 A) Comparative analyses typically investigate evolutionary processes giving rise to trait differences at

685 the tip of the phylogeny, yet these analyses are often limited to species-level (mean) values. B)

686 Biological multilevel meta-analyses typically use highly heterogeneous datasets and partition the

687 variance into different components to explain variation in effect sizes. These analyses also often

688 incorporate publication bias tests and follow rigorous reporting practices. C) The unified approach we
 689 propose merges the strength of both approaches. This approach improves quantifying and
 690 decomposing ecological, methodological, and evolutionary variation in biological datasets.
 691



692
 693 **Figure 2:** Conceptual frameworks for community-level and function-valued analyses. A: Multilevel
 694 comparative analyses can be used to investigate questions at broader scales, by combining results
 695 from spatially dependent models performed at the community level. B: Function-valued analyses use
 696 multivariate analyses to investigate patterns from multiple parameters of a continuous trait (e.g.,
 697 thermal performance curve, TPC). This approach can leverage datasets with missing data, and
 698 investigate overall effects for each parameter, as well as trait covariation and trade-offs. CTmin:
 699 critical thermal minimum; Topt: thermal optimum; CTmax: critical thermal maximum.

700