adiv: An r package to analyse biodiversity in ecology — **Source link** ↗

Sandrine Pavoine

Institutions: Centre national de la recherche scientifique

Published on: 01 Sep 2020 - Methods in Ecology and Evolution (Wiley)

Topics: Ecosystem diversity, Biodiversity and Phylogenetic diversity

Related papers:

- Old and new challenges in using species diversity for assessing biodiversity

- Learning To Measure Biodiversity: Two Agent-Based Models That Simulate Sampling Methods & Provide Data for Calculating Diversity Indices

- Conservation evaluation and phylogenetic diversity

- Protecting Biodiversity (in All Its Complexity): New Models and Methods.

- The dimensionality of biological diversity revisited

adiv: An r package to analyse biodiversity in ecology

Sandrine Pavoine

1 # adiv: an R package to analyse biodiversity in ecology

2 Sandrine Pavoine

3 *Centre d'Ecologie et des Sciences de la Conservation (CESCO), Muséum*

4 *national d'Histoire naturelle (MNHN), Centre National de la Recherche Scientifique (CNRS),*

5 *Sorbonne Université, CP 135, 57 rue Cuvier 75005 Paris, France*

6

7 **Correspondence**

8 Sandrine Pavoine

9 Email: sandrine.pavoine@mnhn.fr

10 Mail: UMR7204 CESCO, MNHN, 61 rue Buffon, 75005 Paris, France

11

12 Running headline: adiv: an R package to analyse biodiversity

13 Article type: Applications

14

15 **Abstract**

16 1. R is an open-source programming environment for statistical computing and graphics

17 structured by numerous contributed packages. The current packages used for biodiversity

18 research focus on limited, particular aspects of biodiversity. Most packages focus on the

19 number and abundance of species.

20 2. I present an R package named `adiv` that provides additional methods to measure and

21 analyse biodiversity. `adiv` contains approaches to quantify species-based, trait-based

22 (functional) and phylogenetic diversity (i) within communities (α diversity), and (ii) between

23 communities (β diversity), and (iii) to partition it over space and time (α, β, and γ levels of

24 diversity). Partitioning approaches allow evaluating whether the levels of α and β diversity

25 could have been obtained by chance. Moreover, groups of biological entities (e.g., species of

26 the same clade or with similar biological characteristics) that drive each level of diversity (α,

27 β, and γ) can be identified via ordination analyses.

28 3. Although the package focuses on interspecific diversity in its current state, the developed

29 approaches can also be applied to analyse intraspecific diversity or, at another level,

30 ecosystem diversity. More generally, the functions can be applied in any discipline interested

31 in the concept of diversity, such as economics or linguistics. Indeed, all available approaches

32 can be easily applied at other scales and to other disciplines provided that the data have the

33 required format: a matrix of abundance or presence/absence data of some entities in some

34 collections and information on the differences between the entities.

35 4. `adiv` aims to complement existing R packages in order to provide scientists with a wide

36 variety of diversity indices, as each index reflects a very specific facet of biodiversity. `adiv`

37 will grow in the future to integrate as many validated approaches for biodiversity analysis as

38 possible, not yet available in R. As it includes both traditional and recent viewpoints on how

39 biodiversity should be evaluated, `adiv` offers a promising platform where methods to analyse

40 biodiversity can be developed and compared in terms of their statistical behaviour and

41 biological relevance. Applications of the most relevant tools for a given study aim will

42 eventually improve research on human-driven variations in biodiversity.

43

44 KEYWORDS: biodiversity, community ecology, conservation, diversity apportionment,

45 diversity measurement, functional trait, phylogeny, software

46

47

1 | INTRODUCTION

The package named `adiv` (for Analysis of bioDIVersity) focuses on the measurement of biological diversity (or biodiversity, the variability in life from genes to ecosystems) and on the analysis of its organization in space and/or time. I developed `adiv` in the open-source R environment for statistical computing and graphics (R Core Team, 2019). Although this environment is rich in contributed packages, most of those that focus on biodiversity, which include approximately 40 packages (among the 15300 current packages, R Core Team, 2019), actually concentrate on *species diversity* (see Table 1 for a glossary with the definitions of the expressions in italics). A few other packages (e.g., `FD` Laliberté, Legendre, & Shipley, 2014; `TPD` Carmona, 2018) focus on *functional diversity*, while others (e.g., `PhyloMeasures` Tsirogiannis & Sandel, 2017) are dedicated to *phylogenetic diversity*. More generally, some packages, e.g., `entropart` (Marcon & Herault, 2015), focus on species, phylogenetic and functional diversity and consider a limited number of diversity indices. Only a few packages focus on other aspects of biodiversity; for example, `diveRsity` (Keenan et al., 2013) evaluates genetic diversity in the context of population genetic analyses.

     `adiv` enriches the total library of biodiversity packages by providing both traditional statistical approaches to diversity (e.g., Hill, 1973) and recent approaches that measure *trait-based* and *phylogenetic diversity* (e.g., Faith, 1992; Chao et al., 2010; Pavoine, Love, Bonsall, 2009; Pavoine, Baguette, Bonsall, 2010; Kondratyeva et al., 2019; Pavoine & Ricotta, 2019). In addition, all approaches can be easily applied to other aspects of diversity. Indeed, `adiv` relies on the following principle: biodiversity emerges from the differences between entities, regardless of the entities selected (e.g., species, individuals, genera or assemblages) and the criterion used to evaluate the differences between the selected entities (e.g., morphology, behaviour, evolution). The main functions of `adiv` permit the quantification of species-based,

trait-based and phylogenetic diversity within communities (α diversity) and between

communities (β diversity). They complement these measurements that partition the diversity

over all communities (γ diversity) with ordination methods to identify groups of biological

entities (e.g., species of the same clade or with similar biological characteristics) that drive

levels of diversity (Figure 1). The package is available on CRAN (https://cran.r-

project.org/web/packages/adiv/index.html). As an R package, it is open-source to encourage

transparency in science, as all codes can be freely read and checked.

2 | DATA SETS

The functions in `adiv` typically work on a matrix of species presence-absence or abundance

data in *communities* and on species trait data and/or phylogenetic (or taxonomic) data (Figure

1). Data that characterize communities can also be included, such as environmental, spatial or

temporal data. As trait-based and phylogenetic aspects of diversity are often compared in

ecological studies, the `adiv` package implements a few functions to test if closely related

species share similar traits (*phylogenetic signal*, functions `rtestdecdiv`, `K`, `Kstar` and `Kw`).

For illustrative purposes, `adiv` currently contains eight data sets. One of these data

sets, named `batcomm`, is used below to illustrate some of functions of `adiv`. `batcomm` is a list

of two components: `batcomm$ab` contains the abundance of 34 bat species in four habitats

(rainforests, cacao plantations, old fields, and cornfields) in the Selva Lacandona of Chiapas,

Mexico (data collected by Medellín, Equihua, & Amin, 2000); `batcomm$tre` contains a

phylogenetic tree. The tree is ultrametric, meaning that the total branch length from any tip

(species) to the root remains constant.

3 | DIVERSITY INDICES

3.1 | Species diversity

`adiv` contains two main functions for *species diversity* indices: `speciesdiv`, which includes widely used indices such as *species richness* and the Shannon (1948) index, and `divparam`, which includes indices that have a parameter to control the importance given to rare versus abundant species in diversity measurements. As an illustration, using `divparam`, I applied Hill (1973) numbers to the bat dataset. Note that all scripts used in this paper are available in the `adiv` vignette (Pavoine, 2020), named "adiv Package User Guide", that can be accessed from the R console by using the following script: `browseVignettes("adiv")`. As a traditional diversity index, the Hill index increases with species richness and evenness in species abundances. Its parameter, denoted q, increases with the importance given to abundant species compared to rare species in diversity measurement (q=0 means species richness, i.e., equal importance to all species). According to the Hill index, the rainforest dominates in terms of species diversity; however, variations among the diversity levels of the other three habitats indicate how the consideration of abundance data may influence our interpretation of the impact of environmental disturbance on biodiversity (Figure 2a).

With species diversity indices, species are implicitly considered interchangeable because only the number and/or abundance of species is important, not their identity. An assemblage of three bird species, say, a blackbird, sparrow and pigeon, would be considered as rich as an assemblage with three species of different groups, say, a blackbird, domestic cat, and common wall lizard. In contrast, by considering phylogenetic analysis rather than strictly presence- or abundance-based analyses, a higher value of diversity could be attributed to the second assemblage than to the first as the species in the second are more phylogenetically distinct and have more divergent biological characteristics or traits.

3.2 | From species diversity to phylogenetic diversity

122   In the `adiv` package, four main functions consider the phylogeny of species when

123   measuring their diversity. Among them, `evodiv` and `evodivparam` consist of replacing

124   species in traditional diversity indices (for `evodiv`) and traditional parametric diversity

125   indices (for `evodivparam`) with branch units or "features" on phylogenetic trees where

126   species are the tips (Faith, 1992; Pavoine & Ricotta, 2019). It is assumed that the number of

127   features supported by a given branch of a phylogenetic tree is equal to the length of that

128   branch, such that the richness of features in a community is Faith's widely used phylogenetic

129   diversity index (Faith, 1992). The abundance of any feature on a given branch of the

130   phylogenetic tree is measured as the summed abundance of all tips (species) descending from

131   that branch. This approach of replacing species by features is simple and can be applied to all

132   diversity indices developed so far. For example, phylogenetic data can be simply added to the

133   analysis of the bat communities using the `evodivparam` function. The results, which are

134   displayed in Figure 2b, show that when abundant species are given high weight ($q > 2$), all

135   habitats reach similarly low levels of phylogenetic diversity, indicating that, within each

136   habitat, abundant species tend to be closely related.

137

138   ## 3.3 | Trait-based diversity

139   The indices dedicated to phylogenetic diversity can be used with trait-based data if a

140   dendrogram is established, for example, by applying a clustering approach to a matrix of trait-

141   based differences between species. Three other functions in `adiv` allow the measurement of

142   diversity using direct trait-based differences or similarities between species: `QE`, `qHdiv` and

143   `Rentropy`. `QE` implements Rao's quadratic entropy (1982), which is the average abundance-

144   weighted trait-based difference between any two species in a community. The more different

145   the traits of any two individuals in a community are, the higher the quadratic entropy is.

146   `Rentropy` is equivalent to `QE` but species' relative abundances are squared-root transformed

6

147   before calculation. Both `QE` and `Rentropy` generalize well-known species diversity indices to

148   include trait-based data: `QE` generalizes the Simpson (1949) index, which is a simple function

149   of the Hill index if its parameter q equals 2, and `Rentropy` is a direct generalization of the Hill

150   index when q=0.5. `qHdiv` allows, when required, intraspecific variation in biological trait

151   values to be considered.

152   These (dis)similarity-based indices can inversely also be used to evaluate phylogenetic

153   diversity if phylogenetic (dis)similarities are used instead of trait-based (dis)similarities. For

154   example, with the bat dataset, I calculated cophenetic distances (sum of branch lengths along

155   the shortest path) between species on the phylogenetic tree and used them in functions `QE` and

156   `Rentropy` (Figure 2c). With `QE`, all habitats had similar levels of phylogenetic diversity,

157   which is in accordance with the results obtained above with function `evodivparam` if

158   parameter q equals 2 (Figure 2b). With `Rentropy`, the rainforest dominates in terms of

159   phylogenetic diversity, with cacao plantations and old fields having intermediate values and

160   cornfields having the lowest value, which is in accordance with the results of function

161   `evodivparam` if q approaches 0.5 (Figure 2b).

162

4 | DISSIMILARITY INDICES

164   The concepts of diversity and (dis)similarity are linked: the diversity of an assemblage is null

165   if all its components are identical. The biodiversity of a region increases with the increase in

166   the dissimilarities between species and also between communities. Indices of dissimilarity

167   thus complement those of diversity in `adiv`.

168

4.1 | From species-to-species to community-to-community dissimilarity

170   indices

171 A few functions are dedicated to the calculation of dissimilarities or similarities between

172 species using trait, taxonomic or phylogenetic data (e.g., `CFprop`, `CFbinary`, `dsimFun`,

173 `dsimTax`, and `dsimTree`). These functions lead to particular mathematical properties for the

174 (dis)similarities between species. These properties named "positive semidefinite" for

175 similarity matrices and "Euclidean" for dissimilarity matrices are exploited, for example, in

176 the `dsimcom` function of `adiv` (see details in the `adiv` vignette named "adiv Package User

177 Guide"). `dsimcom` implements the Pavoine and Ricotta (2014) index of similarity between

178 two communities by comparing their species lists, the abundances of each species and the

179 functional or phylogenetic similarities between the species. This function is restricted to

180 matrices of similarities between species that are said positive semidefinite. Other indices of

181 the (dis)similarity between two communities can integrate any matrix of (dis)similarities

182 between species without any restrictions on their mathematical properties apart from having

183 nonnegative values and sometimes being bounded between 0 and 1. This is the case, for

184 example, for the indices available in the `dissABC` and `dissRicotta` functions of `adiv`.

185

186 4.2 | From compositional dissimilarity to tree-based dissimilarity

187 `adiv` contains two functions dedicated to phylogeny-based indices of the dissimilarity

188 between two communities: `evodiss` and `evodiss_family`. Although these functions are

189 dedicated to use with phylogenetic trees, they can be more generally applied to other tree-

190 shaped data, such as trait-based dendrograms. These functions use the feature perspective

191 described in section "3.2 | From species diversity to phylogenetic diversity", where a feature is

192 a branch unit on a phylogenetic tree. From this perspective, traditional dissimilarity indices

193 are not applied to species presence/absence or abundance data but to the presence/absence or

194 abundance of each feature. This perspective grants access to a family of dissimilarity indices,

195 and the `evodiss` and `evodiss_family` functions contain 30 key indices from this family,

196   including 6 parametric indices where the importance given to rare versus abundant features

197   can be controlled. For example, applying function `evodiss` with the chord distance to the bat

198   communities reveals that the two habitats with the most divergent phylogenetic compositions

199   are the old field and cornfield habitats (phylogenetic dissimilarity=0.30), while the most

200   similar habitats are the rainforest and cacao plantation habitats (0.12). Given the large number

201   of species-based dissimilarity indices developed so far, more indices from this family could

202   be easily added in the future depending on the needs identified in the ecological and

203   conservation literature.

204

5 | APPORTIONMENT AND ORDINATION OF DIVERSITY

5.1 | Alpha, beta, and gamma diversity across space

207   When more than two communities are compared, the dissimilarity among them, named β

208   diversity, complements the diversity within each community, named α diversity. The diversity

209   of the whole set of communities (γ diversity) emerges from the combination of α and β

210   diversity. `adiv` contains functions to partition species-based, trait-based or phylogenetic

211   diversity in a nested hierarchy (with α, β, and γ levels; `abgdivparam`, `abgevodivparam`,

212   `eqRao`, `eqRS`, `eqRSintra`, and `wapqe` functions). For example, `abgevodivparam`

213   implements partitioning of Hill numbers. Applied to the bat dataset, this function shows that

214   phylogenetic differences between the habitats mostly concern rare species, as phylogenetic β

215   diversity decreases with parameter q (Figure 2d).

216   Some partitioning functions are associated with simple permutation tests to evaluate

217   whether each level of diversity could have been obtained by chance. For example, the

218   application of one of these tests (function `rtestEqRS`, permutation test) to the bat dataset

219   shows that the differences in the phylogenetic compositions of the four habitats are not

220   significant when species abundances are considered (statistic of the test, β diversity

9

221 standardized between 0 and 1 = 2.42 10^{-2}, P-value = 0.071), although the differences are

222 significant when only presence/absence data are evaluated (β = 2.37 10^{-2}, P-value = 0.049;

223 nominal α = 0.050). This result confirms that the most abundant species in each habitat tend

224 to be phylogenetically similar, while phylogenetic differences occur for rare species.

225   `adiv` also implements a range of ordination analyses to visualize species and

226 communities as points in a space that reveals which species contribute to the differences

227 between communities according to their traits or taxonomic or phylogenetic positions. For

228 example, the application of *phylogenetic correspondence analysis* (function `evoCA`) to the bat

229 dataset with presence/absence data (Figure 3) shows that cornfields were characterized by the

230 absence of many species and many clades observed in other habitats and the presence of two

231 closely related species: the long-tongued bats *Hylonycteris underwoodi* and *Lichonycteris*

232 *obscura* (Figures 3a-c). However, the representation of the phylogenetic tree on the map of

233 the phylogenetic correspondence analysis highlights that the phylogenetic differences between

234 habitats are small, although significant (see entangled phylogenetic branches on Figure 3d).

235

236 5.2 | Temporal variations and crossed partitioning

237   Although they have thus far been mostly used in the context of spatial hierarchy, all

238 partitioning and ordination approaches cited above could be explored to analyse the temporal

239 variations in species, *taxonomic*, trait-based and phylogenetic *diversity*. Additionally, crossed

240 double principal coordinate analysis (crossed-DPCoA, `crossdpcoa_maineffect`,

241 `crossdpcoa_version1,` and `crossdpcoa_version2` functions) allows for the partitioning of

242 species-based, taxonomic, trait-based or phylogenetic diversity between two crossed factors

243 according to the methods in Pavoine et al. (2013). For example, 10 regions could be each

244 followed for 5 years, resulting in space (regions) and time (years) being two crossed factors

245 affecting diversity. Crossed-DPCoA allows the effects of space to be disentangled from the

246 effects of time with regard to variations in diversity level. `adiv` also allows studying spatial or

247 temporal variations in trait-based diversity in a phylogenetic context thanks to the `decdiv`

248 function. `decdiv` indeed implements the approach by Pavoine, Baguette, and Bonsall (2010)

249 to partition the α, β and γ trait-based diversity across the nodes of a phylogenetic tree (see also

250 `rtest.decdiv` function for associated permutation tests).

251

6 | CONCLUSIONS AND PERSPECTIVES

253 `adiv` focuses on ecological data to analyse species, trait-based and phylogenetic diversity. By

254 extension, it can be used by any discipline interested in the general concept of diversity

255 provided that the formats of the datasets are similar (at least presence/absence or abundance

256 of any entity in any collection of entities and potentially information to characterize the

257 differences and similarities between the entities). `adiv` will grow in the future as a state-of-

258 the-art package for the measurement and statistical analysis of biodiversity. In particular, new

259 developments may concern the following key points:

260 • **Connections with other packages on diversity.** I have written the `adiv` package to

261 complement existing packages on diversity. Some functions contained in other

262 packages (e.g., TPD, Carmona, 2018) are thus not available in `adiv`. Future

263 developments could include the importation of functions from available packages to

264 ease their calculation with the data format used by `adiv`.

265 • **Additions of new methodologies to measure diversity and partition it across**

266 **spatial and temporal units.** Each current R package that tackles the concept of

267 biodiversity contains only a few of the myriad of diversity indices developed so far in

268 the literature. The consequence is that the indices are being used in scientific studies

269 based on their accessibility in popular packages rather than based on their true

270 scientific relevance and interest. Indeed, having these indices in a package renders

them more accessible for researchers for whom programming is a constraint and a

challenge. `adiv` aims to complement these packages to offer ecologists a wider variety

of diversity indices. In the interest of open science in biodiversity research, researchers

around the globe and from any discipline are invited to suggest critical methodologies

for analysing biodiversity that could be implemented in `adiv`.

- **Originality, uniqueness and redundancy.** `adiv` also complements current packages

  by offering ways to evaluate how trait and phylogenetic information impacts

  evaluations of biodiversity. In particular, `adiv` contains functions (`uniqueness` and

  `treeUniqueness`) to quantify how redundant or unique a community is compared to a

  scenario where all species would be maximally dissimilar (e.g., having the most

  distinct values possible for biological traits). It also contains functions (e.g.,

  `distinctDis`, `distinctTree`, `distinctTopo` and `distinctUltra`) to identify

  species that are original in a community because they have unique states of biological

  traits or no close relatives in the phylogeny. The presence of original species increases

  the diversity of the community. Further versions of `adiv` will include more originality

  indices, especially those that account for species abundances and species extinction

  risks (see Kondratyeva, Grandcolas & Pavoine, 2019 and references therein).

Having all these methodologies in the same package will ease the diffusion of statistical

methods among all researchers and data analysts in environmental organizations interested in

the analysis of biodiversity. It will also facilitate comparison among methods in terms of their

statistical behaviour and biological relevance. All these developments in `adiv` will contribute,

over the years, to open science and advances in (bio)diversity research.

ACKNOWLEDGEMENTS

ORCID

Sandrine Pavoine   https://orcid.org/0000-0003-2767-6484

DATA ACCESSIBILITY STATEMENT

All data used in this paper are available online in package `adiv`: https://cran.r-project.org/web/packages/adiv/index.html.

REFERENCES

Carmona, C. P. (2018). TPD: Methods for Measuring Functional Diversity Based on Trait Probability Density. R package version 1.0.0. https://CRAN.R-project.org/package=TPD

Chao, A., Chiu, C. H., & Jost, L. (2010). Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365,* 3599–3609. https://doi.org/10.1098/rstb.2010.0272

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation, 61,* 1–10. https://doi.org/10.1016/0006-3207(92)91201-3

Hill, M. O. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology, 54,* 427–432. https://doi.org/10.2307/1934352

319   Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity:

320   An R package for the estimation of population genetics parameters and their

321   associated errors. *Methods in Ecology and Evolution, 4,* 782–784.

322   https://doi.org/10.1111/2041-210X.12067

323   Kondratyeva, A., Grandcolas, P., & Pavoine, S. (2019). Reconciling the concepts and

324   measures of diversity, rarity and originality in ecology and evolution. *Biological*

325   *Reviews, 94,* 1317–1337. https://doi.org/10.1111/brv.12504

326   Laliberté, E., Legendre, P., & Shipley, B. (2014). FD: measuring functional diversity from

327   multiple traits, and other tools for functional ecology. R package version 1.0-12.

328   https://CRAN.R-project.org/package=FD

329   Marcon, E., & Herault, B. (2015). entropart: An R package to measure and partition diversity.

330   *Journal of Statistical Software, 67,* 1–26. https://10.18637/jss.v067.i08

331   Medellín, R. Equihua M., & Amin, M. A. (2000). Bat diversity and abundance as indicators of

332   disturbance in Neotropical rainforest. *Conservation Biology*, *14,* 1666–1675.

333   https://doi.org/10.1111/j.1523-1739.2000.99068.x

334   Pavoine, S. (2016). A guide through a family of phylogenetic dissimilarity measures among

335   sites. *Oikos, 125,* 1719–1732. https://doi.org/10.1111/oik.03262

336   Pavoine, S. (2020). adiv: Analysis of Diversity. R package version 1.4. https://CRAN.R-

337   project.org/package=adiv

338   Pavoine, S., Baguette, M., & Bonsall, M. B. (2010). Decomposition of trait diversity among

339   the nodes of a phylogenetic tree. *Ecological Monographs, 80,* 485–507.

340   https://doi.org/10.1890/09-1290.1

341   Pavoine, S., Blondel, J., Dufour, A.-B., Gasc, A., & Bonsall, M. B. (2013). A new technique

342   for analysing interacting factors affecting biodiversity patterns: crossed-DPCoA. *PloS*

343   *One, 8,* e54530. https://doi.org/10.1371/journal.pone.0054530

344 Pavoine, S., Love, M., & Bonsall, M. B. (2009). Hierarchical partitioning of evolutionary and

345   ecological patterns in the organization of phylogenetically-structured species

346   assemblages: application to rockfish (genus: Sebastes) in the Southern California

347   Bight. *Ecology Letters, 12,* 898–908. https://doi.org/10.1111/j.1461-

348   0248.2009.01344.x

349 Pavoine, S., & Ricotta, C. (2014). Functional and phylogenetic similarity among

350   communities. *Methods in Ecology and Evolution, 5,* 666–675.

351   https://doi.org/10.1111/2041-210X.12193

352 Pavoine, S., & Ricotta, C. (2019). A simple translation from indices of species diversity to

353   indices of phylogenetic diversity. *Ecological Indicators, 101,* 552–561.

354   https://doi.org/10.1016/j.ecolind.2019.01.052

355 Rao, C. R. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical*

356   *Population Biology, 21,* 24–43. https://doi.org/10.1016/0040-5809(82)90004-1

357 R Core Team (2019). R: A language and environment for statistical computing. R foundation

358   for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

359 Shannon, C. E. (1948). A mathematical theory of communication. *Bell System technical*

360   *journal, 27,* 379–423, 623–656.

361 Simpson, E. H. (1949) Measurement of diversity. *Nature, 163,* 688.

362 Tsirogiannis, C., & Sandel, B. (2017). PhyloMeasures: fast and exact algorithms for

363   computing phylogenetic biodiversity measures. R package version 2.1.

364   https://CRAN.R-project.org/package=PhyloMeasures

365

TABLE 1 Glossary

Concept	Definitions used in this paper
Community	Set of species observed at a given time in a given place.
Functional diversity	Trait-based diversity applied to functional traits (see the definition of trait-based diversity below). A functional trait is a trait that influences the way species (or organisms) respond to environmental conditions or the way they contribute to ecosystem properties.
Phylogenetic diversity	Diversity in the positions of species (or of organisms, for abundance-weighted diversity) on a phylogenetic tree. A community with many closely related species may have less phylogenetic diversity than a community with few distantly related species.
Phylogenetic correspondence analysis	Extension, proposed by Pavoine (2016), of the well-known ordination approach named correspondence analysis to include phylogenetic data.
Phylogenetic signal	Positive correlation between the differences in species trait values and the distances between species on a phylogenetic tree.
Species diversity	Number and abundance of species. Indices of species diversity increase with the number of species in a community and with the evenness in species abundances.
Species evenness	Evenness in species abundances.
Species richness	Number of species.
Taxonomic diversity	Diversity in the taxonomic relationships between species. For example, a community with many species of the same genus

|	may have less taxonomic diversity than a community with few species that belong to different orders.
Trait-based diversity	Diversity in trait values. A community with many organisms that share similar trait values may have lower trait-based diversity than a community with few organisms characterized by distinct trait values.

367

368   FIGURES

369
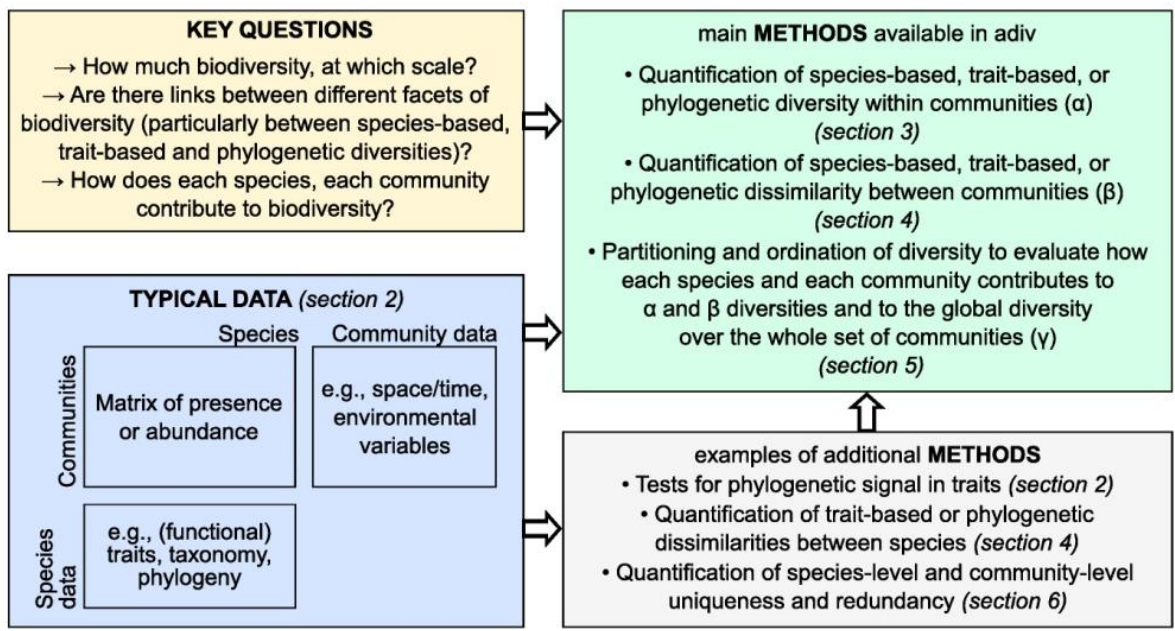
370

371   **FIGURE 1** Schematic organization of `adiv`.

372

373
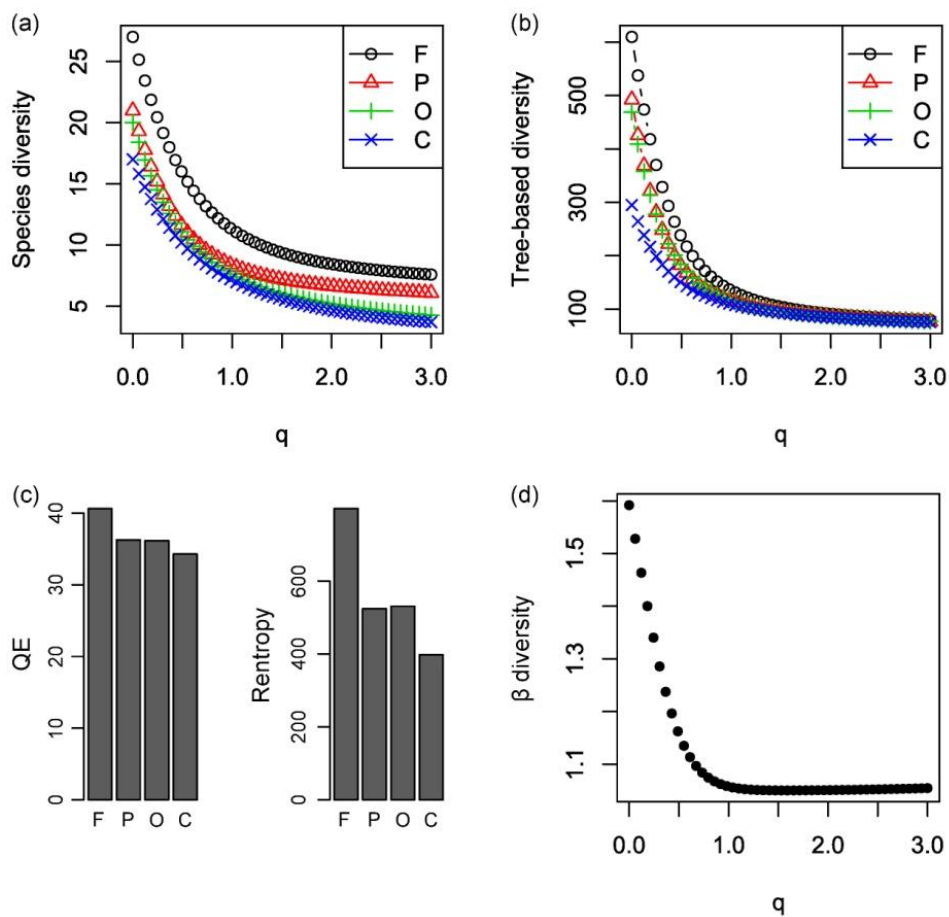
**FIGURE 2** Bat species diversity (a), phylogenetic diversity (b-c) in four habitats

(F=rainforests; P=cacao plantations; O=old fields; C=cornfields), and phylogenetic β

diversity between the habitats (d). In panels (a), (b), and (d), the Hill index is given as

a function of its parameter q. The higher q is, the less importance is given to rare

species (in a) or features (in b and d) in the evaluation of diversity.

379
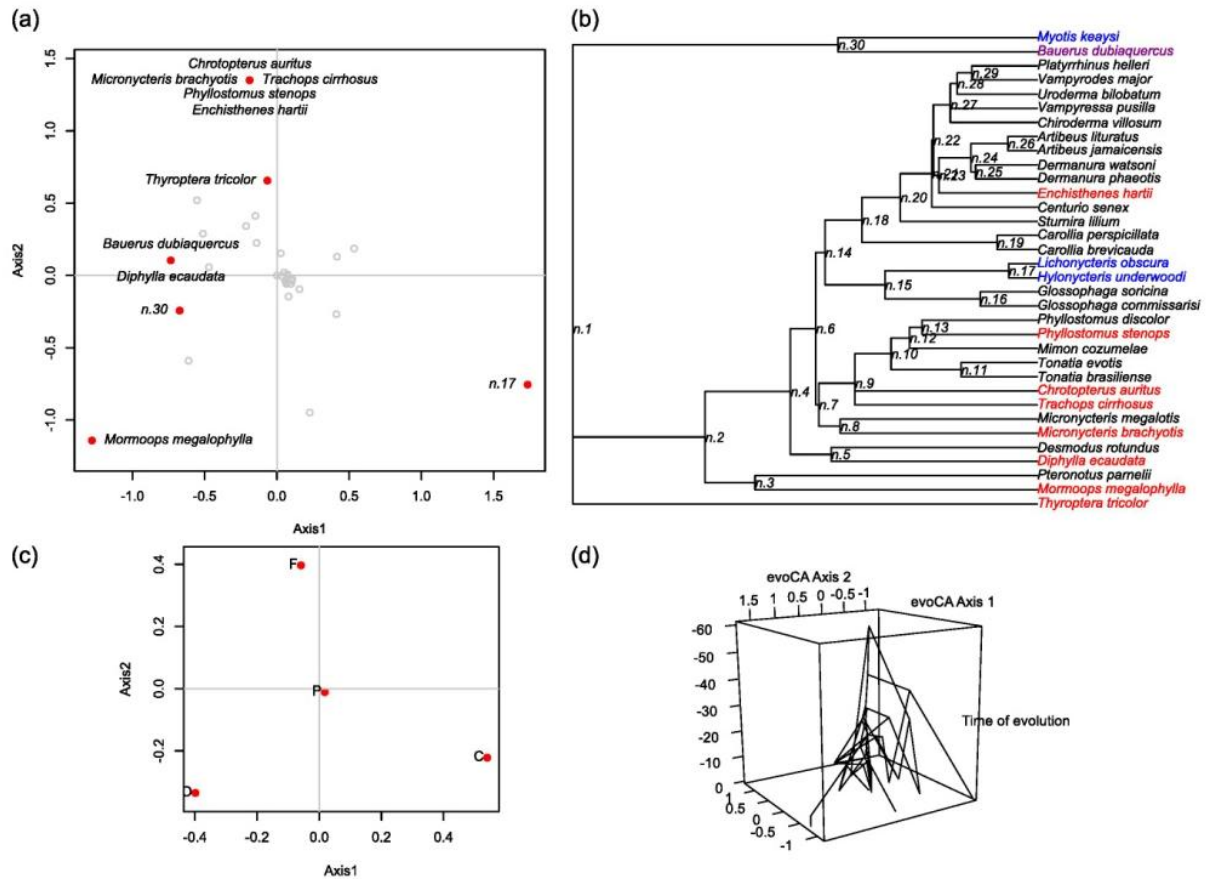
380

381   **FIGURE 3** Analysis of the bat phylogenetic composition of four habitats (F=rainforests;

382   P=cacao plantations; O=old fields; C=cornfields). (a, b, and c) Phylogenetic tree and

383   factorial maps obtained with phylogenetic correspondence analysis (evoCA; the first

384   axis represents 42% of the phylogenetic dissimilarities among habitats, and the second

385   represents 37%). (a) Coordinates for the nodes of the phylogenetic tree; the positions

386   of the "top-contributing" nodes/species (explaining more than 5% of the variance on

387   one of the axes) are highlighted in red and labelled. (b) Phylogenetic tree (in red, top-

388   contributing species; in blue, species descending from the top-contributing nodes; in

389   violet, the top-contributing species that descend from a top-contributing node). (c)

390   Coordinates for the habitats on evoCA axes. (d) 3D map of the phylogenetic tree on

391   evoCA axes (in `adiv`, the plot is interactive).

392