Herodotools: An R package to integrate macroevolution, biogeography, and community ecology

Gabriel Nakamura[1*]; Arthur V Rodrigues[2]; André Luís Luza[2,3]; Renan Maestri[2]; Vanderlei Debastiani[2] and Leandro da Silva Duarte[2]

1- Texas A&M University-Corpus Christi, Department of Life Sciences, 6300 Ocean Dr. 78412

2 – Universidade Federal do Rio Grande do Sul, Departamento de Ecologia, Avenida Bento Gonçalves, Agronomia, Porto Alegre – Rio Grande do Sul, Brasil, 90650-001

3 – Departamento de Ecologia e Evolução. Universidade Federal de Santa Maria, Santa Maria – Rio Grande do Sul

*Corresponding author: gabriel.nakamura.souza@gmail.com

Running title: Integrating macroevolution and ecology

Acknowledgments

We thank Matheus Drumm for developing the early version of functions to calculate assemblage age. GN was supported by CNPq. LD research was supported by a CNPq Productivity Fellowship (grant 307527/2018-2). RM was supported by CNPq (150391/2017-0) and UFRGS. GN and LD are members of the National Institute for Science and Technology (INCT) in Ecology, Evolution, and Biodiversity Conservation, supported by Ministério da Ciência, Tecnologia, Inovações e Comunicações/ CNPq (proc. 465610/2014-5) and FAPEG (proc. 201810267000023). AVR was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, and ALL received post-doctoral

23 fellowships from CNPq (#153024/2022-4, #164240/2021-7, #151228/2021-3, #152410/2020-1).

24 No permits to collect fieldwork data were required in this study.

Abstract

Aim: Historical processes like speciation, extinction, and historical dispersal are the ultimate factors generating and maintaining biodiversity in space and time. While detecting the effect of those processes on the distribution of biodiversity has great relevance by itself, how to measure them is critical to interpreting the underlying causes of biological patterns. However, metrics of macroevolution used at biogeographical scales usually ignore the variation of macroevolutionary processes at scales finer than entire regions. Likewise, biogeography and community ecology often ignore deep-time evolutionary processes, giving us a limited picture of the role of historical processes in community assembly. To overcome this problem, it is necessary to integrate data from ancestral state reconstructions, current species distributions, and biogeographical regionalization. We hereby present Herodotools, an R package that integrates macroevolutionary models with data on the distribution of species occurrences in assemblages and biogeographic regions.

Location: Global application, with an example from Neotropics.

Major taxa studied: Any taxa, with an example from small rodents (genus *Akodon* and subfamily Sigmodontinae).

Methods: We developed an R package called Herodotools, designed to streamline analyses of historical biogeography, including regionalization, calculation of assemblage age, lineage in situ diversification, and community phylogenetic metrics, which merge species occurrence with macroevolutionary methods of ancestral area and trait reconstruction. We described the main functions of our R package through toy examples and illustrated the use of our new package by analyzing the historical biogeography from small rodent assemblages in the Neotropics.

Results: We showed that our methods can integrate methods from biogeography, macroevolution, and community ecology, allowing us to downscale the effects of historical processes and calculate important historical variables (e.g., age of assemblages, in-situ diversification) in different scales, from entire regions to communities of co-occurrent species.

Main Conclusions: Our package provides the first platform to streamline the analysis of historical biogeography, enabling a better understanding of historical processes at different levels of organization, from local assemblages to entire biogeographical regions.

Keywords: Historical biogeography; macroevolutionary dynamics; ancestral state reconstruction; biogeographical regionalization; model-based metrics; diversification; historical dispersal

Introduction

Evolutionary processes such as speciation, extinction, and historical dispersal are considered the ultimate factors promoting the distribution of biological diversity across space and time (Ricklefs, 1987; Ricklefs & Jenkins, 2011; Wiens & Donoghue, 2004a). Classical biogeographical patterns are shaped by those factors jointly with ecological factors (e.g., temperature, pluviosity, food availability). The most famous example is the latitudinal gradient of species richness. Among all the possible explanations for it, at least two of three leading hypotheses consider the potential role of macroevolutionary processes in producing the gradient, namely "time for speciation" (Stephens & Wiens, 2003) and the "diversification rate" hypotheses (Egan et al., 2022; Pontarp et al., 2019). Consequently, a better understanding of ecological patterns depends on using reliable proxies for quantifying macroevolutionary phenomena in biogeographical studies (Houle et al., 2011).

Despite the central role in biogeography, macroevolutionary phenomena have usually been measured by adopting proxies that do not necessarily reflect their nature, especially by ignoring the fine-scale effect of macroevolutionary processes, which are generally accounted for at broad scales (Mouquet et al., 2012). The calculation of evolutionary time (hereafter ET) and diversification metrics illustrates some of these problems. ET is commonly considered an important historical variable by representing the available time lineages had to build up new species in a given environment (Li & Wiens, 2019; Stephens & Wiens, 2003). Therefore, lineages that spend more time in an environment might have more opportunities to accumulate more species, and the differences in ET among areas will drive variation in diversity. ET is usually measured by extracting information about the age of lineages from a phylogenetic tree, for example, by calculating the maximum branch length among species (MBL) (García-Andrade

et al., 2021), or by using ancestral area reconstructions to estimate ancestral colonization times (ACT) for entire biomes (Li & Wiens, 2019). The problem is: while the former equivocally assumes that speciation events at phylogenetic tree are simultaneous with events in geographical space (the age of the oldest species indicates arrival time for the whole assemblage in a given region), the latter adopts the premise that ecological heterogeneity inside a biome is not relevant (e.g., variation in composition associated with core and edge of a biome like showed by Luza et al., 2021) since all the assemblages inside the biome will present the same ACT. Similar problems apply to other metrics often used as proxies for macroevolutionary processes in biogeography, like diversification rates. For example, when comparing two areas with the same species composition using a common tip-based metric of speciation (e.g., DR metric; Jetz et al., 2012), will result in a single value of diversification (or speciation) even if the colonization times differ between them. Therefore, even though diversification is a lineage's property, different areas/assemblages should account for distinct colonization times and time available for diversification, which is usually ignored in biogeographical studies.

Even though common macroevolutionary metrics reflect the nature of macroevolutionary processes under specific scenarios (e.g., García-Andrade et al., 2021) (e.g., high niche conservatism in ancestral area of occurrence), the interaction between ecological variation and macroevolutionary processes are diverse and complex (e.g., Skeels et al., 2022). This fact is an evidence that the use of common proxies for macroevolutionary processes usually does not accurately reflect their nature, causing equivocal interpretations of biogeographical patterns. Consequently, meaningful measures of macroevolutionary processes in biogeographical studies should be able to capture the variation in species occurrence in space and time (e.g., different

rates of diversification in different areas depending on the present-day and past geographical distribution of lineages).

A better way to represent macroevolutionary variables would be using methods that integrate macroevolutionary models of ancestral area and trait reconstruction with tip (species)-based metrics that can be projected at assemblage scale (Mouquet et al., 2012). For example, Van Dijk et al. (2021) coupled a macroevolutionary model of ancestral area reconstruction with species occurrence data to estimate assemblage age and then tested two concurrent macroecological mechanisms of biodiversity assembly ("Out of the tropics" and "Niche conservatism" hypotheses). Another example is the study by Luza et al., (2021) that coupled a macroevolutionary model of trait evolution to estimate tip-based metrics that enabled to understand how diet evolution in Sigmodontinae rodents is shaped by the environment (ecotone and core spatial position). Despite only a few, those methods proved reliable in testing hypotheses that integrate ecological variation at assemblage scale through current species composition and macroevolutionary processes to interpret biogeographical patterns and processes. However, those methods are scattered in different studies and analytical routines hosted in different repositories, with no single platform that allows us to perform those analyses, which limits their usage and, consequently, our ability towards a better understanding of macroevolutionary imprints in biogeographical and community patterns (McGill et al., 2019).

To overcome potential problems of inferring the role of macroevolution in biogeographical and community ecology studies, in this work, we present Herodotools (see https://gabrielnakamura.github.io/Herodotoolshttps://gabrielnakamura.github.io/Herodotools/articles/Intro_Herodotools_vignette for an explanation for the package name and logo), an R package that provides a computational infrastructure that bridges the gap between biogeography,

125 community ecology, and macroevolution. Our new tool overcomes the problems
126 abovementioned by wrapping up in a single platform existing (Luza et al., 2021; Maestri &
127 Duarte, 2020a; Van Dijk et al., 2021) and new methods that provide reliable interpretation of
128 biogeographical patterns and historical processes acting from entire regions to local
129 communities. Here we demonstrate the conceptual rationale used to incorporate macroevolution,
130 biogeography, and community ecology in a single framework (Figure 1). The scheme in Figure 1
131 starts by showing the use of methods that allow detecting the influence of historical processes on
132 specific lineages (Figure 1a represented through ancestral state reconstruction methods, for
133 example, BioGeoBEARS)(Matzke, 2013) and large spatial scales (Figure 1c represented mainly
134 through regionalization methods), mostly ignoring the variation in local communities within a
135 region. On the other hand, community ecology (Figure 1b) deals with local variations of multiple
136 coexisting species. Still, it mostly ignores the effects of macroevolutionary dynamics at this scale
137 or, when considered, using some unreliable proxies of historical imprints based only on present-
138 day community patterns (e.g., community phylogenetic metrics) (Crouch et al., 2019). We fill
139 out these gaps by proposing integrative methods that allow us to evaluate macroevolutionary
140 dynamics in both biogeographic and community scales (Figure 1d). All these methods were
141 implemented in Herodotools package.

142 We aim to show the analytical details behind core package functions and exemplify a
143 broad analytical pipeline to investigate the following questions using a dataset of the Neotropical
144 genus *Akodon* (Maestri et al., 2017, 2019), a species-rich and widespread genus of sigmodontine
145 rodents: i) How to estimate the arrival ages at the assemblage level considering current and past
146 geographical distribution of species and ancestors? ii) What is the relative importance of in-situ
147 diversification and historical dispersal to determine the structure of assemblages? iii) How to

quantify trait evolutionary dynamics at the assemblage scale while considering the variation in assemblage level (this last using the whole Sigmodontinae clade)? These questions represent just a few that can be answered by integrating macroevolution with biogeography and community ecology (McGill et al., 2019) in Herodotools R package.

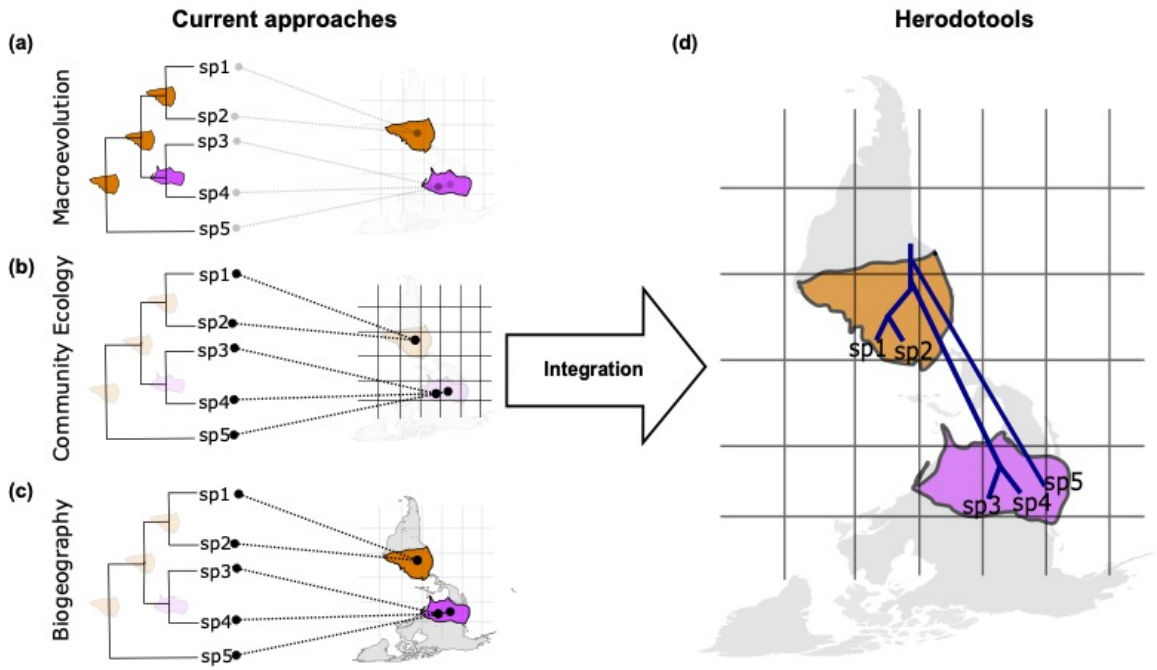


Figure 1: Conceptual representation of current approaches in macroevolution, community ecology, and biogeography (a-c) and their integration in the Herodotools R package (d). Macroevolutionary dynamics (a) usually focus on trait evolution and diversification at biogeographical areas (orange and purple polygons reconstructed across phylogeny nodes) using methods like ancestral state reconstruction (e.g., BioGeoBEARS) but ignoring the variation in local assemblages (grids), that is the domain of community ecology (b) which in its turn mostly

ignores macroevolutionary dynamics in space and/or time. Finally, biogeography investigates large spatial patterns of variation but also disregards the variation of assemblages inside a region (c). Herodotools (d) present functions that fill these gaps by integrating macroevolutionary models in different spatial scales, from local communities to bioregions.[Double column]

Methods

General description of Herodotools package

Herodotools integrate two different data types: one comprising the output from macroevolutionary analysis (*e.g.,* ancestral area and/or trait reconstruction/mapping) and another comprising species occurrences in spatial units such as assemblages or regions (*e.g.,* biomes, ecoregions, evoregions). Specifically, functions implemented in Herodotools allows the manipulation of data from common macroevolutionary analyses, e.g., ancestral area reconstruction models in BioGeoBEARS (Matzke, 2013), and ancestral traits reconstruction (Bollback, 2006), converting the output of these analyses into matrices and tables to calculate macroevolutionary metrics at the assemblage level (Table 1). These metrics can be used to map the effects of historical factors at different scales or as variables in common modeling frameworks, allowing to test hypotheses in ecology, macroevolution, biogeography, and community ecology (e.g., Luza et al., 2021; Van Dijk et al., 2021). Additionally, Herodotools performs phylogenetic regionalization methods, map transition zones (Maestri & Duarte, 2020a), and detects macroevolutionary sources and sinks (Goldberg et al., 2005) (functions '*evoregions'*, '*calc_affiliation_evoreg',* and '*calc_dispersal_from'*, respectively).

181 Table 1: Description of the main functions present in Herodotools package. [Double column]

Fields	Function name	Description	Reference
Data preparation	get_node_range_BioGeoBEARS	Take a BioGeoBEARS output to produce a matrix of ancestral occurrence (assemblages x nodes)	This study
	spp_nodes	Computes a matrix of species (rows) and their respective nodes (columns)	This study
Macroevolution + Biogeography	calc_dispersal_from	Compute the amount of contribution of each ancestral range to the species composition in other regions	This study
Biogeography	evoregions	Computes phylogenetic regionalization based on phylogenetic fuzzy weighted method	Maestri and Duarte (2020)
	calc_affiliation_evoreg	Computes the degree of affiliation of a cell within the region	Maestri and Duarte (2020)
	calc_spp_association_evoreg	Classify species in evoregions	Maestri and Duarte (2020)

	find_max_n_cluster	Computes the maximum number of clusters to be used on '*evoregions'* function	Maestri and Duarte (2020)
Macroevolution + Community Ecology	calc_insitu_metrics	Computes in situ component of ecophylogenetic metrics ($PD_{in\ situ}$ and $PE_{in\ situ}$)	This study
	calc_insitu_diversification	Computes in-situ component of diversification metrics after lineages colonization	This study
	calc_age_arrival	Compute the age of assemblages	Van Dijk et al (2019)
Macroevolution + phenotypic evolution + Community Ecology	calc_tip_based_trait_evo	Computes tip-based metrics that express trait macroevolutionary dynamics (transition rate, last transition time, and stasis time)	Luza et al. (2021)

182

183 The integration of macroevolutionary dynamics into community, biogeographic, and trait
184 analysis comprises two steps: 1) the use of an ancestral area reconstruction model to decompose
185 the evolutionary history dynamics on the phylogenetic tree in two components, namely 'in-situ
186 diversification' and 'ex-situ processes', and 2) use step 1 to calculate tip-based metrics for each
187 lineage in the phylogenetic tree.
188 The in-situ diversification component comprises the evolutionary history that emerged
189 due to in-situ speciation into ecological assemblages. In-situ diversification measures
190 evolutionary history/time from all the events that occurred since each lineage's colonization and
191 establishment in the assemblage's region (i.e., no further dispersal events). In other words, it
192 represents the path from tip to the node linking the current species occurrence in an assemblage
193 and the most recent common ancestor in which the occurrence range was estimated in the same
194 region as the assemblage, estimated through ancestral area reconstruction (Fig. 1d) (Van Dijk et
195 al., 2021). This phylogenetic component is also used to calculate other assemblage level metrics,
196 such as the age of assemblages and the amount of phylogenetic diversity (Faith, 1992) and
197 phylogenetic endemism (Rosauer et al., 2009) that emerge from in-situ diversification into a
198 region (see Table 1 for the list of functions that enable the calculation of these metrics). The ex-
199 situ component corresponds to the evolutionary history that arose due to events of ex-situ
200 diversification and historical dispersal, i.e., events that occurred before the arrival and
201 establishment of a lineage in the assemblage in which the present-day species are occurring.
202 In-situ diversification and ex-situ components are illustrated in Figure 2. This
203 hypothetical example (Fig. 2a) represents a result from an ancestral area reconstruction model,
204 with ancestral regions of occurrence represented by letters A, B, and C. These areas can be
205 interpreted, for example, as biomes (e.g., Maestri et al., 2019; Van Dijk et al., 2021a; Wiens &

Graham, 2005). The matrix in Fig. 2b represents the current area of occurrence for each species in three assemblages (comm 1, comm 2, and comm 3). By recognizing in-situ diversification and ex-situ components, we can decompose the amount of macroevolutionary history that emerged inside a region due to in-situ diversification and the phylogenetic history component from another region (ex-situ). Following this rationale, we can notice that community 1 is assembled only by in-situ diversification (Fig. 2c), community 2 mainly ex-situ events from region C (Fig. 2d), and community 3 is assembled by ex-situ events from region B and A, with a recent contribution from in-situ diversification events (Fig. 2e). The values of in-situ diversification and regular diversification calculated, respectively, accordingly to the metric proposed in this study (in-situ diversification) and the inverse of equal splits metric (DR) (Redding & Mooers, 2006; Jetz et al. 2012) are also shown for each community. These metrics will have the same values when all the diversification occurs inside the region where the community is (community 1). Suppose only part of diversification occurred inside the region where the community is located (communities 2 and 3), in that case, in-situ diversification will be lower than DR. This happens because the in-situ diversification considers the geographical area in which the diversification has occurred.

It is worth noting that Figure 2 is a simplified example of macroevolutionary dynamics, in case of multiple disconnected diversification events (colonization followed by dispersion followed by colonization again), our method will capture only the effects of in-situ diversification after the last colonization, i.e., the effect of colonization after the establishment of a lineage in an area. The numerical example in Figure 2 can be reproduced using Herodotools package examples from function '*calc_insitu_diversification*'.

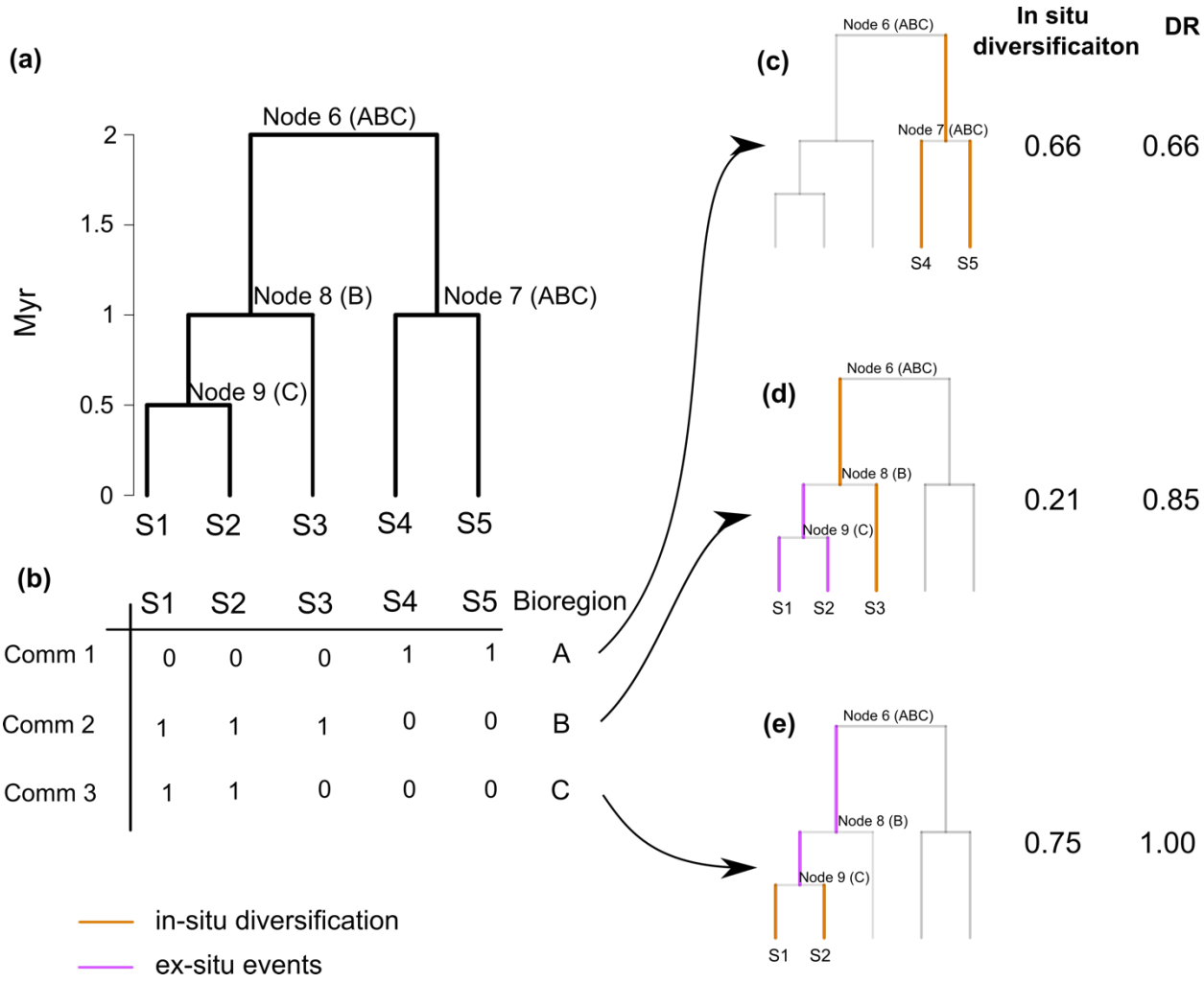


Figure 2: Schematic figure illustrating the decomposition of macroevolutionary dynamics that can be performed in Herodotools. Purple branches in the tree correspond to the evolutionary history that emerged from ex-situ events (e.g., dispersal, speciation), and orange branches emerged from in-situ diversification. (a) A phylogenetic hypothesis with an ancestral area reconstruction (represented by letters in each node). (b) Species incidence in three hypotheticals assemblages and biomes. (c), (d), and (e): Three different hypothetical scenarios of macroevolutionary dynamics for each assemblage with their respective values of in-situ diversification and diversification rate (DR). It is worth noting that in-situ diversification and DR

for different assemblages will have the same values if all the diversification occurs within the assemblages' region. [Single column]

In the next section, we explain in more detail the main functions of the Herodotools package, emphasizing some specific and new functions arising from the integration between macroevolutionary and community data made in this package.

Phylogenetic regionalization and shifts in phylogenetic turnover across bioregions

Methods aiming to define biogeographic regions based on either taxonomic (Edler et al., 2016; Holt et al., 2013; Kreft & Jetz, 2010; Olivero et al., 2013; Vilhena & Antonelli, 2015) or phylogenetic relationships among species of a given biological group (Holt et al., 2013; Maestri & Duarte, 2020a) have been intensively developed over the last decade, using different site resemblance and clustering methods. While all methods are valuable as classification tools for historical biogeography and evolutionary macroecology, bioregions defined from either species composition or the Simpson index of phylogenetic beta diversity (Holt, et al. 2013; Daru et al., 2020) might lead to the detection of evolutionarily unreal biogeographic regions, as regions arising from classifications might lack a coherent, shared history of diversification. It occurs because site resemblance and clustering methods are unable to detect transition zones, i.e., regions where sites show low phylogenetic affinity to their respective biogeographic regions (Maestri & Duarte 2020). On the other hand, classifying biogeographic regions based on evoregions (Maestri & Duarte 2020) enables mapping biogeographic transition zones in addition to core biogeographic regions, better showing intricate species distributions and facilitating the interpretation of biogeographic regions.

As an interesting development, evoregions also allow the interpretation of the historical development of each biogeographic region directly along with the diversification history of a lineage represented as a phylogenetic tree (e.g., Fig 2 in Duarte and Maestri, 2018). Thus, evoregions is a useful methodological approach for historical biogeography and evolutionary macroecology whenever unveiling the geographical history of diversification is a primary goal. Phylogenetic classification with evoregions can be performed using the function '*evoregions'* and detecting phylogenetic turnover zones can be done by using the function '*calc_affiliation_evoreg'* in the Herodotools package.

Metrics for inference of historical processes at assemblage level

One of the main drawbacks in ecology and evolution is the consideration of historical processes at the assemblage level (Mouquet et al., 2012). Herodotools fill this gap by implementing a set of metrics that can be calculated at the assemblage level. The first metric is the age of assemblage, explained in the previous section and calculated with the function '*calc_age_arrival'*. We also implemented tip-based metrics of diversification that account for macroevolutionary history. For example, the function '*calc_insitu_diversification'* modified the commonly used DR metric (Equation 1 DR; Jetz et al., 2012) calculated as the inverse of the mean equal-splits measures (ES; Redding & Mooers, 2006) as follows:

Equation 1 $DR_i = 1/ES_i$

Equation 2 $ES_i = \left(\sum_{j=1}^{N_i} l_j \frac{1}{2^{j-1}} \right)$

In our modification, we first calculate a version of equal-splits measure for each species *i* that considers only the edges *j*, being l_j the length of the edge *j* that emerged after the colonization

(arrival and establishment of lineages in an area up to the present) of species lineage in the regions where the assemblage is placed, what we call $ES_{in\ situ}$ (Equation 2).

Instead of calculating the inverse of ES to obtain the DR for species *i* (Equation 1), we calculate the proportion of $ES_{in\ situ}$ relative to the regular ES (Reeding and Mooers, 2006), and then multiply this value by the total DR (Equation 3).

Equation 3 $In - situ\ diversification = DR_i * \left({ES_{in\ situ}}/{ES}\right)$

With this modification, we obtained an in-situ diversification metric. When the $ES_{in\text{-}situ}$ is low relative to ED, in-situ diversification will be low. When $ES_{in\text{-}situ}$ equals ES, DR, and in-situ diversification will equal since all diversification emerged in-situ for a given lineage after its colonization and establishment. Finally, we calculate in-situ diversification for each assemblage as its harmonic mean across all assemblage species.

We also implemented popular ecophylogenetic metrics, such as Phylogenetic Diversity (PD Faith, 1992) and Phylogenetic Endemism (PE)(Rosauer et al., 2009), that account for in-situ diversification by applying the same rationale. For PD and PE, we modified the original metrics by using only the branch lengths that emerged after the arrival and establishment of the species lineages in an assemblage's region. We then obtained what we called $PD_{in\text{-}situ}$ and $PE_{in\text{-}situ}$.

Mapping trait evolution dynamics over space

As mentioned before, to scale up macroevolution to a macroecological assemblage-based level of analysis, methods should provide species-specific data. A few existent metrics are designed to gather species-specific evolutionary data directly from phylogenies, including estimates of tip-based diversification (Jetz et al., 2012; Redding & Mooers, 2006; Title & Rabosky, 2019), and tip-based trait evolutionary rates (Castiglione et al., 2018). However, these metrics cannot handle

306 temporal variation in trait states and age/time of trait appearance in the history of a
307 clade/phylogeny. To tackle this issue, Luza et al., (2021) formulated an analytical framework
308 that allows analyzing species-specific rates and tempo of (discrete) trait evolution by proposing
309 three new tip-based metrics: i) transition rates, ii) stasis time, and the iii) last transition time.
310 Briefly, these metrics capture the evolutionary history of trait changes from the root to each
311 current species and summarize it in species-specific number of trait state changes (transition
312 rates), the total evolutionary time without change (stasis time), and time since the last change
313 (last transition time). Those metrics can be projected at the assemblage level, for example, by
314 simply averaging species "traits" within assemblages, whereby it is possible to infer ecological
315 and historical processes shaping the rates and tempo of trait evolution in local assemblages.
316 These three tip-based metrics can be calculated with the function '*calc_tip_based_trait_evo'*.
317
318 *Historical biogeography of the Akodon genus*
319 To demonstrate the functionalities of Herodotools, we analyzed a data set of 732 assemblages of
320 the genus *Akodon*. *Akodon* is one of the most species-rich and widely distributed genera of
321 mammals in the Neotropics (Patton et al., 2007, Mammal Diversity Database 2022).
322 Geographically, its 41 described species form two hotspots of richness, one in the Atlantic Forest
323 and the other in the Central Andes, dominating the more inclusive richness pattern of its tribe,
324 the Akodontini (Maestri & Patterson 2016), and overall forming a "dumbell" richness pattern
325 (Pardiñas et al. 2015) also due to its absence in the Amazon. Such bimodal richness peaks and
326 the phylogenetic distribution of its species cast doubt on the geographic origins of the genus,
327 with hypotheses over the years lending support for either an Andean or an Atlantic center of

origination and main diversification of the inclusive tribe (Reig, 1987; D'Elía & Pardiñas 2015; Maestri et al. 2019).

To calculate the importance of in-situ diversification, historical dispersal events and estimate the age of assemblages, we first applied a phylogenetic regionalization method based on evolutionary turnover (Maestri & Duarte, 2020a) implemented in the function '*evoregion'* of Herodotools package. Based on the groups generated by the phylogenetic regionalization, we estimated species' ancestral range using BioGeoBEARS (Matzke, 2013). We built six different models implemented in BioGeoBEARS: DIVA, DEC, and BayArea; each with and without a jump parameter. Details of model construction and the code used can be found in the online resource (https://gabrielnakamura.github.io/Herodotools/articles/Intro_Herodotools_vignette.html). We allowed species to belong to up to three biomes. We performed a model selection using Akaike Information Criterion (AIC) to select the best model for ancestral range estimates. Three models (DEC, DEC+J, and BayArea) presented $\Delta AIC < 2$ and were considered equivalent. We chose the DEC model for further analysis because it has the lowest AIC value and fewer model parameters. The selected model was then used in Herodotools to calculate metrics that represent historical processes at assemblage level. Specifically, we calculated: 1) the age of each assemblage as the mean age across all lineages in an assemblage in which the ancestors of each present-day species arrived and established in the region of a given cell (function '*calc_age_arrival'*); 2) In-situ diversification, obtained by calculating the tip-based metrics of diversification for each species as being the inverse of equal splits metric (Jetz et al., 2012; Redding & Mooers, 2006), but now considering only the branches of the lineage that emerged from an in-situ diversification process (function '*calc_insitu_diversification',* Eq. 2); 3) The contribution of historical dispersal events

for each assemblage, represented as the percentage of species that dispersed from a focal ancestral range for all other regions. The assemblages comprised 1x1 degree grids cells (110 x 110 km around the Equator).

Finally, we analyzed the macroevolutionary dynamics of traits by calculating transition rates, stasis time, and the last transition time for species foraging strata of 214 sigmodontine rodent species (Luza et al., 2021). The results of trait macroevolutionary dynamics are shown only in the online and pdf supplement (https://gabrielnakamura.github.io/Herodotools/articles/Intro_Herodotools_vignette.html), as well as other examples illustrating the use of additional functions implemented in Herodotools package.

Results

Regionalization with evoregion

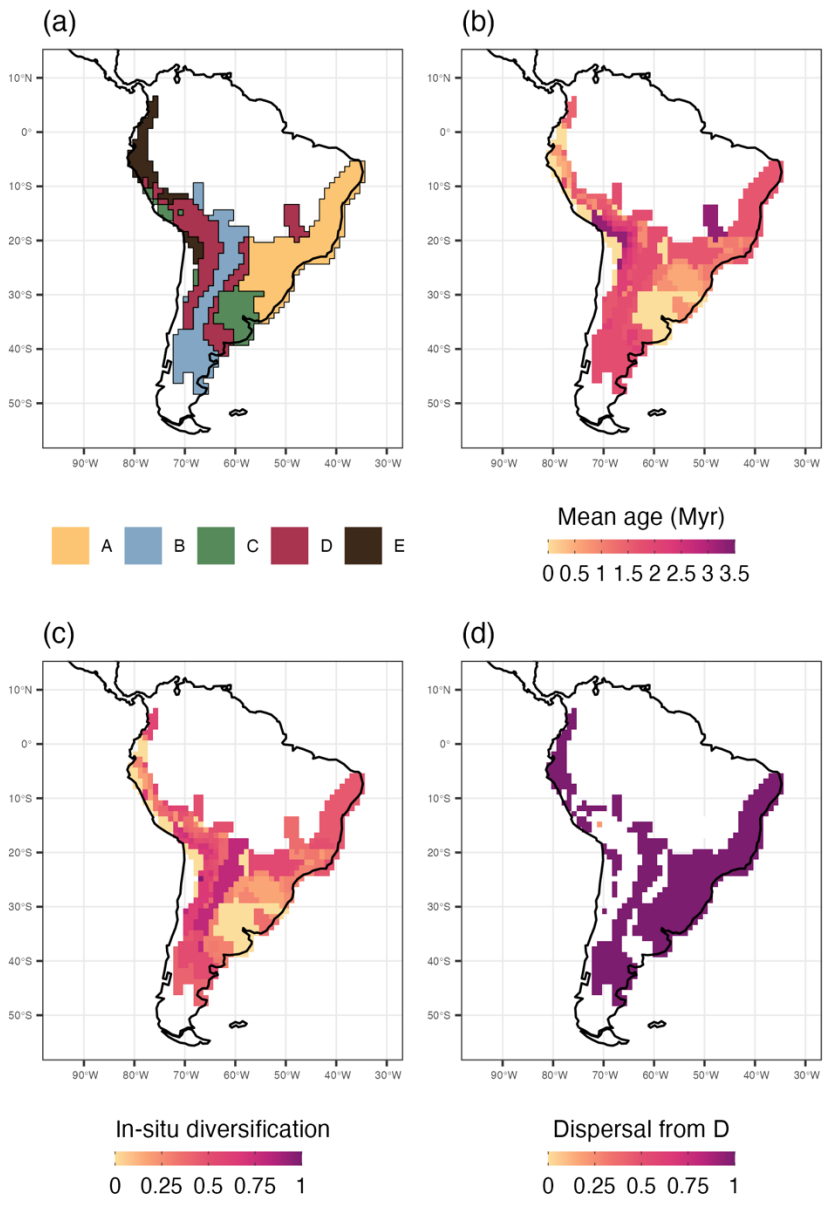
In our empirical example using the genus *Akodon*, we detected five distinct regions using the evoregions approach which we named evoregions A-E (Figure 3a). Evoregion D captures a mix of species from the four species complexes within *Akodon*: *boliviensis*, *cursor*, *aerosus*, and *dolores*. We found components of all these complexes within Evoregion D, culminating in its central position where the richness peak is located, plus idiosyncrasies such as the occurrence of *A. lindberghi*, a species whose species group is not easily defined (Gonçalves et al. 2007; Jayat et al. 2010; Coyner et al. 2013). Members of the cursor group clearly defined Evoregion A; evoregion C was mainly related to the group *boliviensis* plus *A. azarae*; evoregion E was primarily determined by members of the *aerosus* group, and evoregion B by members of the *dolores* group. Furthermore, regarding the affiliation of assemblages to each evoregion, we could

observe that the lowest values of affiliation were found close to the boundaries of evoregions. Furthermore, evoregion D and the south portion of evoregion A presented the lowest affiliation values (i.e., when a cell has a low chance of belonging to the region in which it was classified), indicating zones of high phylogenetic turnover and multiple colonization events (Figure 3 in Supplementary online and pdf material).

Assemblage level metrics – Age and in-situ diversification

Our estimates of assemblage age indicated that assemblages did not present high variation in the age in which ancestors arrived and colonized the assemblages, except for evoregions C and E, which showed the most recent assemblages. On the other hand, ancient assemblages are within regions B and D (around 2 – 2.5 million years since colonization and establishment of lineages in assemblages) (Fig. 3b).

The diversification (DR) and in-situ diversification metrics showed similar spatial patterns (Figure 5 in online and pdf supplementary material https://gabrielnakamura.github.io/Herodotools/articles/Intro_Herodotools_vignette.html), with higher values of in-situ diversification in the evoregion B, with some assemblages presenting almost all the diversification occurring inside this region (Fig. 3c). On the other hand, evoregion C was the one presenting assemblages with the lowest values of in-situ diversification. Together, the age of assemblages and in-situ diversification patterns result from an explosive diversification in a few million years experienced by Akodontini rodents.



394

395 Figure 3: Spatial representation of evoregions (a) and historical variables (b-d). (b) represents the
396 age of assemblages, (c) the in-situ diversification as a proportion of the total diversification
397 (calculated as the DR metric), and (d) represents the proportion of the contribution of region D
398 with lineages for all other evoregions. [Single column]

399

Historical dispersal patterns

Our analysis of historical dispersal showed that evoregion D was the region that most contributed to lineage dispersal for other evoregions. Assemblages in evoregions A, B, C, and E were almost entirely constituted of lineages from evoregion D (Fig. 3d). On the other hand, evoregions A and B had only a small contribution to lineage dispersal. The contribution of different regions to lineage dispersal events is shown in the supplementary online material, Figure 6 in the online vignette (https://gabrielnakamura.github.io/Herodotools/articles/Intro_Herodotools_vignette.html and pdf supplementary material). Results on trait evolution metrics (Transition Rates, Stasis Time, and Last Transition Time) in sigmodontine rodent assemblages can also be found in Supplementary material, Figure 8 (https://gabrielnakamura.github.io/Herodotools/articles/Intro_Herodotools_vignette.html#tip-based-metrics-of-trait-evolution and pdf supplementary material). Those metrics evidenced that sigmodontide assemblages are highly dynamic, with intermediate transition rates, short stasis time over their history, and recent (~2 my) last transitions to current foraging strata. Together, these results indicate that transitions among foraging strata (below ground, ground, trees) were frequent during the evolutionary history of this group in different environments.

Discussion

Herodotools provides a step forward in biogeography studies by wrapping up a set of metrics and methods (Luza et al., 2021; Maestri & Duarte, 2020a; Van Dijk et al., 2021) that better reflect the nature of historical processes in shaping biogeographical patterns at different scales. The methods implemented in our package differ from commonly used metrics in biogeography, mainly by offering tools for accounting for the effects of the historical process on different

aspects of macroevolution and biodiversity patterns in assemblages. This aspect overcomes the issue of analyzing the effect of the macroevolutionary process predominantly at broader scales or single lineages. With Herodotools, the effects of macroevolutionary phenomena can be projected at the community scale (e.g., age and diversification measurements that vary depending on the scale or from one community to the other, even inside the same region) and, for example, can be correlated with ecological variables. Furthermore, we also provide some new versions of classic metrics of phylogenetic diversity and endemism ($PD_{in\text{-}situ}$ and $PE_{in\text{-}situ}$) that incorporate essential components of the evolutionary history of clades at the assemblage level (in situ diversification, $PD_{in\ situ,}$ and $PE_{in\ situ}$), allowing us to understand the role of diversification in generating and maintenance of biodiversity patterns. Therefore, with Herodotools, we provide a single platform that advances the measurement of key variables used in hypothesis testing and the tools that facilitate the calculation and reproduction of those metrics in biogeographical studies.

Despite some of the metrics presented here are not new in biogeographical literature (Luza et al., 2021; Maestri & Duarte, 2020a; Van Dijk et al., 2021), by wrapping them into a single platform, we provide an easy and unified tool to investigate questions in historical biogeography. As far as we know, the Herodotools R package is the first computational infrastructure of analysis that allows considering historical variables (macroevolution) at community and assemblage levels. This integration is essential in improving our understanding of phylogenetic patterns in ecological contexts (Mouquet et al., 2012).

Other approaches that aim to provide macroevolutionary measurements to interpret ecological and biogeographical patterns are available, but with some different applications, making Herodotools a complementary tool for studies in the interface of macroevolution, biogeography, and community ecology. For example, DAMOCLES (Pigot & Etienne, 2009),

446 also allows investigating the macroevolutionary dynamics at the community level. However, it
447 relies on a null model that incorporates historical processes for testing whether there is non-
448 randomness in community phylogenetic structure. The methods presented in Herodotools differ
449 from DAMOCLES since the former aims to decompose the phylogenetic metrics in components
450 that indicate two opposite processes, in situ diversification and ex-situ historical events (among
451 them historical dispersal). In contrast, the last is focused more on hypothesis testing to unveil the
452 mechanisms responsible for the phylogenetic structure of communities.
453 Another significant functionality in Herodotools is the estimation of assemblage age and
454 diversification metrics that properly integrate ancestral and present-day distribution with spatial
455 variation represented by biogeographic regionalization (e.g., ecoregions, evoregions). Estimates
456 of age from Herodotools (Van Dijk et al., 2021) differ from previous studies (García-Andrade et
457 al., 2021; García-Rodríguez et al., 2021; Wiens et al., 2011, 2011; Wiens & Donoghue, 2004b)
458 by allowing the calculation of age for each lineage (tip-based), considering the arrival times of
459 each lineage in different biogeographical regions, and projection at assemblage scale. This means
460 that colonization times can be different between two regions even if both regions are composed
461 of the same species, which is more reliable than the other methods that assign a single
462 colonization time for a species throughout its spatial distribution (Li & Wiens, 2019; Wiens et
463 al., 2011). Also, age values can vary considerably depending on the species composition of
464 assemblages within a given region. This characteristic also applies to other metrics of
465 Herodotools, such as the in-situ diversification (function '*calc_insitu_diversification'*), $PD_{in\ situ}$,
466 and $PE_{in\ situ}$ (function '*calc_insitu_metrics'*). Besides being more reliable by considering the
467 effects of macroevolutionary processes on finer scales (community), the metrics of Herodotools

can be used in common frameworks of hypothesis testing as, for example, linear models relating age and diversity or diversification and richness from biogeographical to assemblage level.

Regarding some further studies in which Herodotools can be used. First is to investigate how different age estimates explain/relate to biogeographical patterns (Wiens et al., 2011; Wiens & Graham, 2005). Despite the many ways to quantify the age of assemblages, up to now, there is no consensual metric, and Herodotools brings a new way to quantify age by projecting ancestral area reconstruction models output at community/assemblage scales. So, it is important to investigate how these different age measurements impact the interpretation of macroecological and macroevolutionary patterns. Another potential application of Herodotools in macroevolutionary studies is regarding the investigation of the relationship between colonization of new areas triggering fast diversification in lineages (ecological opportunity hypothesis) (Burbrink & Pyron, 2009).

Regarding our empirical example, by using a phylogenetic regionalization scheme (Maestri & Duarte, 2020b) for the genus *Akodon,* we were able to depict the uneven geographic distribution of its internal monophyletic groups according to the phylogeny used here (Maestri et al., 2017; Upham et al., 2019), which also closely match other phylogenetic reconstructions. We could also trace back the dispersal history of this lineage, which showed that dispersal from evoregion D to others was found to be the most prominent compared to dispersal from other evoregions (more details on historical dispersal can be found in Supplementary material https://gabrielnakamura.github.io/Herodotools/articles/Intro_Herodotools_vignette.html#historical-dispersal-events-1). The high contribution of region D to historical dispersal events to other areas, together with the low values of affiliation of this region, indicates that the lineages occupying the geographical space of evoregion D present more shifts on phylogenetic

composition over time than the other regions. Overall, the empirical analysis illustrates that Herodotools can be used as a standard pipeline for historical biogeographical analysis and to obtain metrics that reliably represent macroevolutionary processes in different assemblages and biogeographical contexts (arrival ages, in-situ diversification, and dispersal).

Future improvements include implementing functions integrating macroevolutionary models from other popular programs like RevBayes (Landis et al., 2013), and also models of continuous trait evolution to calculate macroevolutionary trait dynamics in the assemblage (e.g., Castiglione et al., 2018). Also, further investigation is needed regarding different ways to calculate in-situ diversification. In the current format, we proposed a metric that only considers in-situ events after the last colonization and the establishment of each lineage in the focal community/region. This implies that multiple in-situ diversification events separated by dispersal events are not fully considered in our metrics. Therefore, the interpretation for our metric of in-situ diversification must be in terms of the amount of in-situ diversification in each assemblage after colonization and the establishment of lineages up to the present. Despite acknowledging this limitation, we argue here that our metrics are still a reliable way to investigate historical imprints of diversification more straightforwardly since we consider a direct estimate of colonization time rather than using proxies of time available for diversification commonly used in community ecology (e.g., community phylogenetic metrics or MBL).

We also plan improvements in a future version of the package to make the integration among macroevolutionary models and assemblage data easier and straightforward in a way that Herodotools can work as the primary toolkit for researchers working in the interface of macroevolution, biogeography, and community ecology.

Data availability statement

All data used in this work is publicly available at https://github.com/GabrielNakamura/Herodotools/tree/main/inst/extdata. The code used to analyze *Akodon* and sigmodontinae assemblages is available in the vignette file (Intro_Herodotools_vignette.Rmd), and the code used to produce Figure 3 is also available at the end of the vignette file, but not shown at the online and pdf supplementary material. This code can be accessed by opening the source file (Intro_Herodotools_vignette.Rmd).

References

Bollback, J. P. (2006). SIMMAP: Stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics*, *7*(1), 88. https://doi.org/10.1186/1471-2105-7-88

Burbrink, F. T., & Pyron, R. A. (2009). HOW DOES ECOLOGICAL OPPORTUNITY INFLUENCE RATES OF SPECIATION, EXTINCTION, AND MORPHOLOGICAL DIVERSIFICATION IN NEW WORLD RATSNAKES (TRIBE LAMPROPELTINI)?: ECOLOGICAL OPPORTUNITY AND DIVERSIFICATION IN SNAKES. *Evolution*, *64*(4), 934–943. https://doi.org/10.1111/j.1558-5646.2009.00888.x

Castiglione, S., Tesone, G., Piccolo, M., Melchionna, M., Mondanaro, A., Serio, C., Di Febbraro, M., & Raia, P. (2018). A new method for testing evolutionary rate variation and shifts in phenotypic evolution. *Methods in Ecology and Evolution*, *9*(4), 974–983. https://doi.org/10.1111/2041-210X.12954

Crouch, N. M. A., Capurucho, J. M. G., Hackett, S. J., & Bates, J. M. (2019). Evaluating the contribution of dispersal to community structure in Neotropical passerine birds. *Ecography*, *42*(2), 390–399. https://doi.org/10.1111/ecog.03927

Daru, B. H., Karunarathne, P., & Schliep, K. (2020). phyloregion: R package for biogeographical regionalization and macroecology. *Methods in Ecology and Evolution*, *11*(11), 1483–1491. https://doi.org/10.1111/2041-210X.13478

Edler, D., Guedes, T., Zizka, A., Rosvall, M., & Antonelli, A. (2016). Infomap Bioregions: Interactive Mapping of Biogeographical Regions from Species Distributions. *Systematic Biology*, syw087. https://doi.org/10.1093/sysbio/syw087

Egan, J. P., Bloom, D. D., & Simons, A. M. (2022). Time for speciation and niche conservatism explain the latitudinal diversity gradient in clupeiform fishes. *Journal of Biogeography*, jbi.14465. https://doi.org/10.1111/jbi.14465

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

García-Andrade, A. B., Carvajal-Quintero, J. D., Tedesco, P. A., & Villalobos, F. (2021). Evolutionary and environmental drivers of species richness in poeciliid fishes across the Americas. *Global Ecology and Biogeography*, *30*(6), 1245–1257. https://doi.org/10.1111/geb.13299

García-Rodríguez, A., Velasco, J. A., Villalobos, F., & Parra-Olea, G. (2021). Effects of evolutionary time, speciation rates and local abiotic conditions on the origin and maintenance of amphibian montane diversity. *Global Ecology and Biogeography*, *30*(3), 674–684. https://doi.org/10.1111/geb.13249

Goldberg, E. E., Roy, K., Lande, R., & Jablonski, D. (2005). Diversity, Endemism, and Age Distributions in Macroevolutionary Sources and Sinks. *The American Naturalist*, *165*(6), 623–633. https://doi.org/10.1086/430012

Holt, B. G., Lessard, J.-P., Borregaard, M. K., Fritz, S. A., Araújo, M. B., Dimitrov, D., Fabre, P.-H., Graham, C. H., Graves, G. R., Jønsson, K. A., Nogués-Bravo, D., Wang, Z., Whittaker, R. J., Fjeldså, J., & Rahbek, C. (2013). An Update of Wallace's Zoogeographic Regions of the World. *Science*, *339*(6115), 74–78. https://doi.org/10.1126/science.1228282

Houle, D., Pélabon, C., Wagner, G. P., & Hansen, T. F. (2011). Measurement and Meaning in Biology. *The Quarterly Review of Biology*, *86*(1), 3–34. https://doi.org/10.1086/658408

565 Isaac, N. J. B., Turvey, S. T., Collen, B., Waterman, C., & Baillie, J. E. M. (2007). Mammals on
566 the EDGE: Conservation priorities based on threat and phylogeny. *PLoS ONE*, *2*(3).
567 https://doi.org/10.1371/journal.pone.0000296
568 Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity
569 of birds in space and time. *Nature*, *491*(7424), 444–448.
570 https://doi.org/10.1038/nature11631
571 Kreft, H., & Jetz, W. (2010). A framework for delineating biogeographical regions based on
572 species distributions: Global quantitative biogeographical regionalizations. *Journal of*
573 *Biogeography*, *37*(11), 2029–2053. https://doi.org/10.1111/j.1365-2699.2010.02375.x
574 Landis, M. J., Matzke, N. J., Moore, B. R., & Huelsenbeck, J. P. (2013). Bayesian analysis of
575 biogeography when the number of areas is large. *Systematic Biology*, *62*(6), 789–804.
576 https://doi.org/10.1093/sysbio/syt040
577 Li, H., & Wiens, J. J. (2019). Time Explains Regional Richness Patterns within Clades More
578 Often than Diversification Rates or Area. *The American Naturalist*, *193*(4), 514–529.
579 https://doi.org/10.1086/702253
580 Luza, A. L., Maestri, R., Debastiani, V. J., Patterson, B. D., Maria, S., & Leandro, H. (2021). Is
581 evolution faster at ecotones? A test using rates and tempo of diet transitions in
582 Neotropical Sigmodontinae (Rodentia , Cricetidae). *Ecology and Evolution*, *December*,
583 18676–18690. https://doi.org/10.1002/ece3.8476
584 Maestri, R., & Duarte, L. D. S. (2020a). Evoregions: Mapping shifts in phylogenetic turnover
585 across biogeographic regions Renan Maestri. *Methods in Ecology and Evolution*,
586 *2020*(August), 1652–1662. https://doi.org/10.1111/2041-210X.13492

Maestri, R., & Duarte, L. D. S. (2020b). Evoregions: Mapping shifts in phylogenetic turnover across biogeographic regions Renan Maestri. *Methods in Ecology and Evolution*, *2020*(August), 1652–1662. https://doi.org/10.1111/2041-210X.13492

Maestri, R., Monteiro, L. R., Fornel, R., Upham, N. S., Patterson, B. D., & de Freitas, T. R. O. (2017). The ecology of a continental evolutionary radiation: Is the radiation of sigmodontine rodents adaptive? *Evolution*, *71*(3), 610–632. https://doi.org/10.1111/evo.13155

Maestri, R., Upham, N. S., & Patterson, B. D. (2019). Tracing the diversification history of a Neogene rodent invasion into South America. *Ecography*, *42*(4), 683–695. https://doi.org/10.1111/ecog.04102

Matzke, N. J. (2013). Probabilistic historical biogeography: New models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography*, *5*(4). https://doi.org/10.21425/f5fbg19694

McGill, B. J., Chase, J. M., Hortal, J., Overcast, I., Rominger, A. J., Rosindell, J., Borges, P. A. V., Emerson, B. C., Etienne, R. S., Hickerson, M. J., Mahler, D. L., Massol, F., McGaughran, A., Neves, P., Parent, C., Patiño, J., Ruffley, M., Wagner, C. E., & Gillespie, R. (2019). Unifying macroecology and macroevolution to answer fundamental questions about biodiversity. *Global Ecology and Biogeography*, *28*(12), 1925–1936. https://doi.org/10.1111/geb.13020

Mouquet, N., Devictor, V., Meynard, C. N., Munoz, F., Couteron, P., Dalecky, A., Gravel, D., Hardy, O. J., Jabot, F., Prinzing, A., Rodrigues, A. S. L., Rohr, R. P., & Th, E. (2012). Ecophylogenetics: Advances and perspectives. *Biological Reviews*, *87*, 769–785. https://doi.org/10.1111/j.1469-185X.2012.00224.x

Olivero, J., Márquez, A. L., & Real, R. (2013). Integrating Fuzzy Logic and Statistics to Improve the Reliable Delimitation of Biogeographic Regions and Transition Zones. *Systematic Biology*, *62*(1), 1–21. https://doi.org/10.1093/sysbio/sys061

Patton, J. L., Pardiñas, U. F. J., & D'Eli?a, G. (2007). *Mammals of South America*. University of Chicago press.

Pontarp, M., Bunnefeld, L., Cabral, J. S., Etienne, R. S., Fritz, S. A., Gillespie, R., Graham, C. H., Hagen, O., Hartig, F., Huang, S., Jansson, R., Maliet, O., Münkemüller, T., Pellissier, L., Rangel, T. F., Storch, D., Wiegand, T., & Hurlbert, A. H. (2019). The Latitudinal Diversity Gradient: Novel Understanding through Mechanistic Eco-evolutionary Models. *Trends in Ecology and Evolution*, *34*(3), 211–223. https://doi.org/10.1016/j.tree.2018.11.009

Redding, D. W., & Mooers, A. O. (2006). Incorporating evolutionary measures into conservation prioritization. *Conservation Biology*, *20*(6), 1670–1678. https://doi.org/10.1111/j.1523-1739.2006.00555.x

Ricklefs, R. E. (1987). Community Diversity: Relative Roles of Local and Regional Processes. *Science*, *235*(4785), 167–171. https://doi.org/10.1126/science.235.4785.167

Ricklefs, R. E., & Jenkins, D. G. (2011). Biogeography and ecology: Towards the integration of two disciplines. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1576), 2438–2448. https://doi.org/10.1098/rstb.2011.0066

Rosauer, D., Laffan, S. W., Crisp, M. D., Donnellan, S. C., & Cook, L. G. (2009). Phylogenetic endemism: A new approach for identifying geographical concentrations of evolutionary history. *Molecular Ecology*, *18*(19), 4061–4072. https://doi.org/10.1111/j.1365-294X.2009.04311.x

633 Skeels, A., Bach, W., Hagen, O., Jetz, W., & Pellissier, L. (2022). Temperature-Dependent
634 Evolutionary Speed Shapes the Evolution of Biodiversity Patterns Across Tetrapod
635 Radiations. *Systematic Biology*, syac048. https://doi.org/10.1093/sysbio/syac048
636 Stephens, P. R., & Wiens, J. J. (2003). Explaining Species Richness from Continents to
637 Communities: The Time-for-Speciation Effect in Emydid Turtles. *The American*
638 *Naturalist*, *161*(1), 112–128. https://doi.org/10.1086/345091
639 Title, P. O., & Rabosky, D. L. (2019). Tip rates, phylogenies and diversification: What are we
640 estimating, and how good are the estimates? *Methods in Ecology and Evolution*, *10*(6),
641 821–834. https://doi.org/10.1111/2041-210X.13153
642 Upham, N. S., Esselstyn, J. A., & Jetz, W. (2019). Inferring the mammal tree: Species-level sets
643 of phylogenies for questions in ecology, evolution, and conservation. *PLOS Biology*,
644 *17*(12), e3000494. https://doi.org/10.1371/journal.pbio.3000494
645 Van Dijk, A., Nakamura, G., Rodrigues, A. V., Maestri, R., & Duarte, L. (2021). Imprints of
646 tropical niche conservatism and historical dispersal in the radiation of Tyrannidae (Aves:
647 Passeriformes). *Biological Journal of the Linnean Society*, *134*(1), 57–67.
648 https://doi.org/10.1093/biolinnean/blab079
649 Vilhena, D. A., & Antonelli, A. (2015). A network approach for identifying and delimiting
650 biogeographical regions. *Nature Communications*, *6*.
651 https://doi.org/10.1038/ncomms7848
652 Wiens, J. J., & Donoghue, M. J. (2004a). Historical biogeography, ecology and species richness.
653 *Trends in Ecology and Evolution*, *19*(12), 639–644.
654 https://doi.org/10.1016/j.tree.2004.09.011

Wiens, J. J., & Donoghue, M. J. (2004b). Historical biogeography, ecology and species richness. *Trends in Ecology and Evolution*, *19*(12), 639–644. https://doi.org/10.1016/j.tree.2004.09.011

Wiens, J. J., & Graham, C. H. (2005). Niche Conservatism: Integrating Evolution, Ecology, and Conservation Biology. *Annual Review of Ecology, Evolution, and Systematics*, *36*(36), 519–539. https://doi.org/10.1146/annurev.ecolsys.36.102803.095431

Wiens, J. J., Pyron, R. A., & Moen, D. S. (2011). Phylogenetic origins of local-scale diversity patterns and the causes of Amazonian megadiversity: Phylogeny and local richness. *Ecology Letters*, *14*(7), 643–652. https://doi.org/10.1111/j.1461-0248.2011.01625.x

Biosketch

Gabriel Nakamura is an ecologist interested in developing numerical and statistical tools to unveil biodiversity patterns and processes, mainly at community and macroecological scales. All the other authors are interested in questions on the interface of ecology, evolution, and historical biogeography. This paper is a result of a collective effort from many years of discussions and informal meetings at the Laboratory of Phylogenetic and Functional Ecology (LEFF) led by Professor Leandro Duarte.

Author contributions

GN build the package and wrote the initial drafts of the manuscript. GN, ALL, and AVR performed all analyses. GN and LDSD conceived the study. ALL and AVR provided substantial input to the R package. VD provided early versions of functions used in the package. RM provided empirical data for analysis. All authors contributed reviewing all draft versions of the manuscript and with at least one function to the package.