1 gen3sis: the general engine for eco-evolutionary 2 simulations on the origins of biodiversity

3 Oskar Hagen^{1, 2 *}, Benjamin Flück^{1, 2}, Fabian Fopp^{1, 2}, Juliano S. Cabral³, Florian Hartig⁴,
4 Mikael Pontarp⁵, Thiago F. Rangel⁶, Loïc Pellissier^{1, 2}

5 ¹ Landscape Ecology, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH
6 Zürich, Zürich, Switzerland

7 ² Land Change Science Research Unit, Swiss Federal Institute for Forest, Snow and Landscape Research, WSL,
8 Birmensdorf, Switzerland

9 ³ Ecosystem Modeling, Center for Computational and Theoretical Biology (CCTB), University of Würzburg, Clara-
10 Oppenheimer-Weg 32, 97074 Würzburg, Germany

11 ⁴ Theoretical Ecology, University of Regensburg, Regensburg, Germany

12 ⁵ Department of Biology, Lund University, Biology Building, Sölvegatan 35, 223 62 Lund, Sweden

13 ⁶ Department of Ecology, Institute of Biological Sciences, Federal University of Goiás, Goiânia Brazil

14 * Correspondence to oskar@hagen.bio

Abstract

Understanding the origins of biodiversity has been an aspiration since the days of early naturalists. The immense complexity of ecological, evolutionary and spatial processes, however, has made this goal elusive to this day. Computer models serve progress in many scientific fields, but in the fields of macroecology and macroevolution, eco-evolutionary models are comparatively less developed. We present a general, spatially-explicit, eco-evolutionary engine with a modular implementation that enables the modelling of multiple macroecological and macroevolutionary processes and feedbacks across representative spatio-temporally dynamic landscapes. Modelled processes can include environmental filtering, biotic interactions, dispersal, speciation and evolution of ecological traits. Commonly observed biodiversity patterns, such as α , β and γ diversity, species ranges, ecological traits and phylogenies, emerge as simulations proceed. As a case study, we examined alternative hypotheses expected to have shaped the latitudinal diversity gradient (LDG) during the Earth's Cenozoic era. We found that a carrying capacity linked with energy was the only model variant that could simultaneously produce a realistic LDG, species range size frequencies, and phylogenetic tree balance. The model engine is open source and available as an R-package, enabling future exploration of various landscapes and biological processes, while outputs can be linked with a variety of empirical biodiversity patterns. This work represents a step towards a numeric and mechanistic understanding of the physical and biological processes that shape Earth's biodiversity.

Keywords

biodiversity, eco-evolution, latitudinal diversity gradient, mechanistic model, modelling framework

38 Introduction

39 Ecological and evolutionary processes have created various patterns of diversity in living
40 organisms across the globe [1]. Species richness varies across regions, such as continents
41 [2, 3], and along spatial and environmental gradients [4, 5], such as latitude [6, 7]. These well-
42 known patterns, derived from the observed multitude of life forms on Earth, have intrigued
43 naturalists for centuries [1, 8, 9] and stimulated the formulation of numerous hypotheses to
44 explain their origin [e.g. 1, 6, 7, 10, 11-15]. Ecologists and evolutionary biologists have
45 attempted to test and disentangle these hypotheses [16], for example via models of
46 cladogenesis [17] or correlative spatial analyses [18, 19]. However, to this day, a mechanistic
47 understanding of ecological, evolutionary and geodynamical spatial dynamics driving diversity
48 patterns remains elusive [20, 21].

49 The complexity of interacting ecological, evolutionary and spatial processes limits our ability
50 to formulate, test and apply the mechanisms forming biodiversity patterns [22, 23].
51 Additionally, multiple processes act and interact with different relative strengths across spatio-
52 temporal scales [20]. Current research suggests that allopatric [24-26] and ecological [22]
53 speciation, dispersal [27] and adaptation [28] all act conjointly in interaction with the
54 environment [29, 30], producing observed biodiversity patterns [31]. Comprehensive
55 explanations of the origin and dynamics of biodiversity must therefore consider a large number
56 of biological processes and feedbacks [32], including species' ecological and evolutionary
57 responses to their dynamic abiotic environment, acting on both ecological and evolutionary
58 time scales [20, 33]. Consequently, biodiversity patterns can rarely be explained by a single
59 hypothesis, as the expectations of the various contending mechanisms are not clearly
60 asserted [20, 34].

61 A decade ago, a seminal paper by Gotelli and colleagues [35] formulated the goal of
62 developing a “general simulation model for macroecology and macroevolution” (hereafter
63 computer models). Since then, many authors have reiterated this call for a broader use of

gen3sis: engine for eco-evolutionary biodiversity modelling

computer models in biodiversity research [20, 36, 37]. With computer models, researchers can explore with simulations the implications of implemented hypotheses and mechanisms and evaluate whether emerging simulated patterns are compatible with observations. Several case studies have illustrated the feasibility and usefulness of computer models in guiding intuition for the interpretation of empirical data [24, 26, 38-42]. Moreover, models have reproduced realistic large-scale biodiversity patterns, such as along latitude [25, 43, 44], by considering climate and geological dynamics [24, 26, 42], and population isolation by considering dispersal ability and geographic distance [24-26, 38-42]. For example, computer models were used to examine how oceans' paleogeography influenced biodiversity dynamics in marine ecosystems [24, 41-43]. Nevertheless, the potential of computer models to enlighten the mechanisms underlying biodiversity patterns remains largely untapped.

Macroevolutionary studies have highlighted that patterns emerging from simulations are generally sensitive to the mechanisms implemented, and to the landscapes upon which those act [24, 25, 42, 43]. Systematically comparing and exploring the effects of mechanisms and landscapes, however, is often hindered by the lack of flexibility and idiosyncrasies of existing models. Most models implement, and thus test, only a limited set of evolutionary processes and hypotheses. Many models are designed for specific and therefore fixed purposes including spatial and temporal boundaries, ranging from the global [24, 25] to continental [26] or regional scale [39, 40], and in time, from millions of years [39, 40, 42, 43] to thousands of years [25, 26]. Moreover, previous eco-evolutionary population models were developed to test a fixed number of mechanisms [24, 25, 35, 40, 42, 44-50]. The diverse input and output formats and limited code availability [51], as well as the different algorithmic implementations, have reduced interoperability between hitherto available models. Biological hypotheses and landscapes should be compared within a common and standardized platform with the modularity required for flexible explorations of multiple landscapes and processes [35]. Increased generality is thus a desirable feature of computer models that aim to explore the mechanisms and landscapes that shape biodiversity in dynamic systems such as rivers [52],

gen3sis: engine for eco-evolutionary biodiversity modelling

oceans [41, 42], islands [39, 40, 53] and mountains [54, 55], or across gradients such as latitude [20, 25, 43].

Here, we present a modelling engine that offers the possibility to explore eco-evolutionary dynamics of lineages under a broad range of biological processes and landscapes. Simulated species populations occupy a spatial domain (hereafter site) bounded by a combination of geological, climatic and ecological factors. The sites occupied by a species define the species' realized geographic range (hereafter species range) [56]. The engine then tracks species populations over time, which can change as a result of dynamic environments, as well as species dispersal ability, ecological interactions, local adaptation and speciation. The initial species range and the criteria for speciation, dispersal, ecological interactions and trait evolution are adjustable mechanisms, allowing the integration of a wide range of hypotheses within the model. Given the flexibility of modifying both mechanisms and landscapes, we consider the engine a general tool and named it "general engine for eco-evolutionary simulations" (hereafter gen3sis). We highlight the potential of gen3sis as a flexible tool to gain inferences about the underlying processes behind biodiversity patterns by tackling a long-standing topic in evolutionary ecology: the latitudinal diversity gradient (LDG) [20]. We implement three alternative hypotheses proposed to explain the LDG [20]: (i) *time for species accumulation* [57-60], (ii) *diversification rates* i.e. depending on temperature [61, 62], and *ecological limits* i.e. depending on energetic carrying capacity [63, 64]. We compare simulation results to empirical distribution and phylogenetic patterns of major tetrapod clades (i.e. mammals, birds, amphibians and reptiles).

Engine principles and scope

Gen3sis is a modelling engine, developed for formalizing and testing multiple hypotheses about the emergence of biodiversity patterns. The engine simulates the consequences of multiple customizable processes and landscapes responsible for the appearance (speciation) and disappearance (extinction) of species over evolutionary time scales. Speciation and extinction emerge from ecological and evolutionary mechanisms dependent on dispersal, species interactions, trait evolution and geographic isolation processes. Customizable eco-evolutionary processes, which interact with dynamic landscapes, make it possible to adjust for various macro-eco-evolutionary hypotheses about specific taxonomic groups, ecosystem types or processes. We made the engine openly available to the research community in an R-package to catalyse an interdisciplinary exploration, application and quantification of the mechanisms behind biodiversity dynamics. The R statistical programming language and environment [65] is widely used for reproducible and open-source research, and since its origins it has been used for handling and analysing spatial data [66]. Gen3sis follows best practices for scientific computing [67], including high modularization; consistent naming, style and formatting; single and meaningful authoritative representation; automated workflows; version control; continuous integration; and extensive documentation.

Gen3sis operates over a grid-based landscape, either the entire globe or a specific region. The landscape used as input is defined by the shape of the colonizable habitat (e.g. land masses for terrestrial organisms), its environmental properties (e.g. temperature and aridity) and its connectivity to dispersal (e.g. the influence of barriers, such as rivers and oceans for terrestrial organisms). Gen3sis simulates species' population range dynamics, traits, diversification and spatial biodiversity patterns in response to geological, biological and environmental drivers. Using a combined trait-based and biological species concept, gen3sis tracks the creation, dynamics and extinction of species ranges, which are composed by a set of sites occupied by species populations. Eco-evolutionary dynamics are driven by user-

gen3sis: engine for eco-evolutionary biodiversity modelling

specified landscapes and processes, including ecology, dispersal, speciation and evolution (Figure 1). Below we explain the gen3sis inputs, the configurations (including eco-evolutionary processes), and the landscapes defining the computer model, as well as user-defined outputs (Figure 1 C–F)

Inputs and initialization

Gen3sis has two input objects which define a particular model (Figure 1). These inputs are: (i) a dynamic landscape (Figure 1 A), which is further divided into environmental variables and distance matrices; and (ii) a configuration (Figure 1 B), in which the user can define initial conditions, biological functions and their parameter values, as well as technical settings for the model core.

Figure 1. Schematic of the main components of the computer model: (A, B) model inputs, including the spatio-temporal landscape objects and the configuration file; (C–F) model outputs, including present and past species ranges, phylogenetic relationships among species, and the ecological traits of species; (G) model engine containing the mechanics; and (H) empirical data applicable for model validation.

154 *Landscape*

155 The landscape objects (Figure 1 A) form the spatio-temporal context in which the processes
 156 of speciation, dispersal, evolution and ecology take place. Landscape objects are generated
 157 based on temporal sequences of landscapes in the form of raster files, which are summarized
 158 in the form of two classes. The first landscape class contains: (i) the geographic coordinates
 159 of the landscape sites, (ii) the corresponding information on which sites are generally suitable
 160 for a clade (e.g. land or ocean), and (iii) the environmental conditions (e.g. temperature and
 161 aridity). The landscape may be simplified into a single geographic axis [e.g. 68] for theoretical
 162 experiments, or it may consider realistic configurations aimed at reproducing real local or
 163 global landscapes [24, 69, 70]. The second landscape class defines the connectivity of the
 164 landscape, used for computing dispersal and consequently isolation of populations. By default,
 165 the connection cost between occupied sites is computed for each time-step from the gridded
 166 landscape data based on haversine geographic distances. This can be modified by a user-
 167 defined cost function in order to account for barriers with different strengths (e.g. based on
 168 elevation [69], water or land) or even to facilitate dispersal in specific directions (e.g. to account
 169 for currents and river flow directions). The final connection costs are stored as sparse distance
 170 matrices [71]. Distance matrices, containing the connection costs, are provided at every time-
 171 step as either: (i) a pre-computed full distance matrix, containing all habitable sites in the
 172 landscape (faster simulations but more storage required); or (ii) a local distance matrix,
 173 computed from neighbouring site distances up to a user-defined range limit (slower
 174 simulations but less storage required).

175 Configuration

176 The configuration object (Figure 1 B) includes the customizable *initialization*, *observer*,
 177 *speciation*, *dispersal*, *evolution* and *ecology* functions. These six functions define a
 178 configuration applied in the simulation engine (Table 1). The possibility to customize these
 179 functions confers the high flexibility of gen3sis in terms of including a wide range of
 180 mechanisms, as illustrated by three configurations presented in a case study (Note S1, Table
 181 S1). Additionally, the configuration object lists the model settings, including: (i) whether a
 182 random seed is used, allowing simulation reproducibility; (ii) start and end times of the
 183 simulation; (iii) rules about aborting the simulation, including the maximum global or local
 184 species number allowed; and (iv) the list of ecological traits considered in the simulation. One
 185 or multiple traits can be defined, which should correspond to those used in the *ecology*
 186 function. Moreover, the *initialization* function creates the ancestor species at the start of the
 187 simulation. Users can define the number of ancestor species, their distribution within the
 188 ancient landscape and their initial trait values. With the *observer* function, changes over time
 189 in any abiotic or biotic information of the virtual world can be recorded by defining the outputs
 190 that are saved at specified time-steps. Outputs can be saved and plotted in real-time as the
 191 model runs. The core biological functions (i.e. *speciation*, *dispersal*, *evolution* and *ecology*)
 192 are presented below.

193 *Core functions and objects*

194 The states of the computer model are updated in discrete time-steps. At each time-step, the
 195 *speciation*, *dispersal*, *evolution* and *ecology* functions are executed sequentially (Figure 2).
 196 Speciation and extinction emerge from interactions across core functions. For example,
 197 speciation events are influenced by *speciation* function as well as by the *ecology* and *dispersal*
 198 functions that interact in a dynamic landscape, ultimately dictating populations' geographic
 199 isolation. Likewise, global extinctions depend on local extinctions, which are influenced by the
 200 *dispersal*, *evolution* and *ecology* functions that dictate adaptation and migration capacity.
 201 Internally, the computer model defines core objects of the simulations: species abundances;
 202 species trait values; the species divergence matrix between all populations for each species;
 203 and the phylogeny of all species created during the simulation. In the following sections, we
 204 describe the core processes in gen3sis, as well as their inputs and outputs. For a summary
 205 see Table 1.

206 Running a simulation in gen3sis consists of the following steps: (i) Read in the configuration
 207 object, prepare the output directories, load the initial landscape (Figure 2 A) and create the
 208 ancestor specie(s) (using the *initialization* function *create_ancestor_species*). (ii) Run the main
 209 loop over the landscape time-steps. At every time-step, the engine loads the appropriate
 210 landscape, removes all sites that became uninhabitable in the new time-step, and executes
 211 the core functions as defined by the configuration object (Figure 2 B). (iii) At the end of every
 212 time-step, gen3sis saves the species richness, genealogy and, if desired, the species,
 213 landscape and other customized observations that are defined in the *observer* function (e.g.
 214 summary statistics and species pattern plots). Core functions are modifiable and can account
 215 for a wide range of mechanisms, as illustrated in the case study (Notes S1 and S2).
 216 Conversely, functions can be turned off, for example in an ecologically neutral model. For a
 217 pseudo-code of gen3sis see Note S3.

gen3sis: engine for eco-evolutionary biodiversity modelling

A Landscape

B Core processes

Figure 2. Schematic example of the *gen3sis* engine simulation cycle of one species' populations over a landscape evolution example containing highlands (yellow), lowlands (green) and a river acting as a barrier (blue). (A) Landscape. A time series of landscapes is used as input, with the landscape being updated after every time-step of the simulation cycle, i.e. after the ecology process. (B) Model core processes. First, the speciation process determines the divergence between geographic clusters of populations that are not connected and splits the clusters into new species if a threshold is reached. In this illustration, divergence between clusters of fox populations was not sufficient to trigger speciation. Second, in the dispersal process, the species spreads within a landscape to reachable new sites. In this illustration, the river limits dispersal. Third, the evolution process can modify the value of the traits in the populations. In this illustration, two fox populations show trait evolution in their ability to cope with the local environment (i.e. red and white fox populations). Fourth, the ecology process recalculates the abundance of the species in each site based on the abiotic condition and co-occurring species, possibly resulting in local extinctions. In this illustration,

gen3sis: engine for eco-evolutionary biodiversity modelling

233 the red fox was unsuited to the lowlands while the white fox survived in the highlands.
234 Speciation and extinction events emerge from multiple simulation cycles of customizable
235 processes.

236 **Table 1.** Presentation of the core functions of *speciation*, *dispersal*, *ecology* and *evolution*
237 implemented in gen3sis. The computation of core functions is customizable in the
238 configuration object. Shown are input objects that are combined to generate updated
239 outputs. The table corresponds to the mechanisms presented in Figure 2 B.

	Objective	Input	Computation	Output
Speciation				
	Determines the divergence between geographic clusters of populations within a species; determines cladogenesis.	Species divergence matrix; species trait matrix; species abundance matrix; landscape values; distance matrix.	Divergence between geographically isolated clusters of populations increases over time while (re-)connected clusters decrease down to zero; speciation happens when the divergence between two clusters is above the speciation threshold, but can also consider trait differences.	Updated species divergence matrix; new species if speciation occurred; updated genealogy table.
Dispersal				
	Determines the colonization of vacant sites.	Species trait matrix; species abundance matrix; landscape values; distance matrix.	Species disperse according to a unique value or a distribution of dispersal values.	Updated species abundance matrix.
Evolution				
	Determines the change of species traits in each site, anagenesis.	Species trait matrix; species abundance matrix; landscape values; geographic clusters; distance matrix.	Traits might change for each species in the populations of occupied sites.	Updated species trait matrix.
Ecology				
	Determines the species abundance in each site.	Species trait matrix; species abundance matrix; landscape values; genealogy.	Change the species abundance, based on landscape environmental values and species co-occurrences; changes species trait values.	Updated species abundance matrix.

Speciation

Core. The *speciation* function iterates over every species separately, registers populations' geographic occupancy (species range), and determines when geographic isolation between population clusters is sufficient to trigger a lineage-splitting event of cladogenesis. A species' range can be segregated into spatially discontinuous geographic clusters of sites and is determined by multiple other processes. The clustering of occupied sites is based on the species' dispersal capacity and the landscape connection costs. Over time, disconnected clusters gradually accumulate incompatibility (divergence), analogous to genetic differentiation. Disconnected species population clusters that maintain geographic isolation for a prolonged period of time will result in different species after the differentiation threshold τ is reached (modelling Dobzhansky-Muller incompatibilities [72]). These clusters become two or more distinct species, and a divergence matrix reset follows. On the other hand, if geographic clusters come into secondary contact before the speciation occurs, they coalesce and incompatibilities are gradually reduced to zero.

Non-exhaustive modification possibilities. A customizable *speciation* function can further embrace processes that modulate speciation. Increased divergence values per time-step can be constant for all species or change depending on biotic and abiotic conditions, such as faster divergence between species occupying higher temperature sites [62], or they can be dependent on population size [73] or other attributes [74]. The function also takes the ecological traits as input, thus allowing for ecological speciation [22], where speciation depends on the divergence of ecological traits between – but not within – clusters [75].

Dispersal

Core. The *dispersal* function iterates over all species populations and determines the connectivity between sites and the colonization of new sites in the grid cell. Dispersal distances are drawn following a user-defined dispersal function and then compared with the distance between pairs of occupied and unoccupied sites. A unique dispersal value can be

gen3sis: engine for eco-evolutionary biodiversity modelling

used (deterministic connection of sites) or dispersal values can be selected from a specified distribution (stochastic connection of sites). If the occupied to unoccupied site connection cost is lower than the dispersal distance, the dispersal is successful. If populations in multiple origin sites manage to colonize an unoccupied site, a colonizer is selected randomly to seed the traits for the newly occupied site.

Non-exhaustive modification possibilities. A customizable *dispersal* function enables the modelling of different dispersal kernels depending on the type of organism considered. Dispersal values can be further linked with: the *ecology* function, for instance trade-off with other traits [76], e.g. dispersal versus competitive ability [77]; and the *evolution* function allowing dispersal to evolve, resulting in species with different dispersal abilities [78].

Evolution

Core. The *evolution* function determines the change in the traits of each population in occupied sites of each species. Traits are defined in the configuration object and can evolve over time for each species' populations. The function iterates over every population of a species and modifies the trait(s) according to the specified function. Any number of traits, informed at the configuration object, can evolve (e.g. traits related to dispersal, niche or competition).

Non-exhaustive modification possibilities. A customizable *evolution* function takes as input the species abundance, species trait, species divergence clusters and the landscape values. In the function it is possible to define which traits evolve and how they change at each time-step. In particular, the frequency and/or amount of change can be made dependent on temperature [79], ecological traits [80], or abundances [81], while the directions of change can follow local optima or various evolutionary models, including Brownian motion [82] and Ornstein–Uhlenbeck [83].

gen3sis: engine for eco-evolutionary biodiversity modelling

290 Ecology

291 **Core.** The *ecology* function determines the abundance or presence of populations in occupied
 292 sites of each species. Thus, extinction processes derive from *ecology* function interactions
 293 with other processes and landscape dynamics. The function iterates over all occupied sites
 294 and updates the species population abundances or presences on the basis of local
 295 environmental values, updated co-occurrence patterns and species traits.

296 **Non-exhaustive modification possibilities.** A customizable *ecology* function takes as input
 297 the species abundance, species trait, species divergence and clusters, and the landscape
 298 values. Inspired by classic niche theory [10, 15, 84], the function can account for various niche
 299 mechanisms, from simple environmental limits to complex multi-species interactions. It is
 300 possible, for example, to include a carrying capacity for the total number of individuals or
 301 species [21] or competition between species based on phylogenetic or trait distances [26],
 302 based on an interaction currency [85], or further constrained by a functional trade-off [76].

303

304 *Outputs and comparisons with empirical data*

305 The computer model delivers a wide range of outputs that can be compared with empirical
 306 data (Figure 1, Table 2). Gen3sis is therefore suitable for analysing the links between
 307 interacting processes and their multidimensional emergent patterns. By recording the time and
 308 origin of all speciation events, as well as trait distributions and abundance throughout
 309 evolutionary history, the simulation model records the information required to track the
 310 dynamics of diversity and the shaping of phylogenetic relationships. The most common
 311 patterns observed and studied by ecologists and evolutionary biologists, including species
 312 ranges, abundances and richness, are emergent properties of the modelled processes (Table
 313 2). All internal objects are accessible to the observer function, which is configurable and
 314 executed during simulation runs. This provides direct simulation outputs in a format ready to
 315 be stored, analysed and compared with empirical data. Given the flexibility of gen3sis, it is
 316 possible to explore not only parameter ranges guided by prior knowledge available for a given
 317 taxonomic group, but also variations in landscape scenarios and mechanisms (Figure 3).
 318 Furthermore, validating modelled outputs with multiple empirical patterns is recommended
 319 [20, 23, 35]. Gen3sis generates multiple outputs, which can be compared with empirical data
 320 using simulation rankings or acceptance criteria [23, 35, 86].

321

322

gen3sis: engine for eco-evolutionary biodiversity modelling

Table 2. List of outputs from the gen3sis computer model, both direct and indirect, that can be compared with empirical data. Direct outputs are the species abundance matrix, species trait matrix and phylogeny, while indirect outputs result from various combinations of the direct outputs. The computations of indirect outputs rely on other packages available in the R environment [65].

Pattern		Scale					
		Spatial			Temporal		
		-		+	-		+
Metric	Example	local	regional	global	present	past	deep past
Alpha diversity (α)	Local species richness follows marked spatial gradients, such as along latitude (LDG, Ricklefs in [87]). Species richness is further correlated across scales when the regional species pool size is positively associated with local species richness [e.g. 4, 88].	*	*	*	*	*	*
Beta diversity (β)	Species turnover is marked along both spatial and environmental gradients [89, 90] and can display sharp boundaries forming biogeographic domains [91].		*	*	*	*	*
Gamma diversity (γ)	Regional difference in species richness, for instance across biogeographic regions with comparable climates, such as the continental temperate region of North America versus Asia [92].			*	*	*	*
Species abundance, frequency and range	Assemblages are generally composed of a few very abundant species and many rare species [93, 94]. A few species tend to occupy many sites, while most are very rare and have a restricted range size [95].	*	*	*	*	*	*
Species ecological niche width distribution	Niche width is heterogeneous across species [96, 97], and narrow niche width leads to higher speciation [98].	*	*	*	*	*	*
Trait evolutionary rates	Ecological traits and niches generally evolve slowly so that closely related lineages have similar traits and niches, coined as niche conservatism [58].	*	*	*	*	*	*
Species diversification rates	Species diversification rate varies over time and across clades [99-101].	*	*	*		*	*
Topological and temporal phylogenetic properties	Empirical phylogenetic trees typically display a topological signature [102] and have more divided branching over time, with marked prevalence of a recent branching distribution [103].	*	*	*		*	*
Phylogenetic alpha (α) and beta (β) diversity	Local communities can show either phylogenetic over-dispersion or clustering compared with the regional pool [104]; greater geographic distances correspond to increased phylogenetic β diversity across biogeographic barriers [105]; decay in phylogenetic similarity with increasing geographic distance [106].	*	*	*		*	*
Functional alpha (α) and beta (β) diversity	Local assemblages represent a subset of the regional functional diversity; functional traits show a typical turnover spatially, often along environmental gradients [107].	*	*	*	*	*	*

Case study: The emergence of the LDG from environmental changes of the Cenozoic.

Context

The LDG is one of Earth's most iconic biodiversity patterns, but the underlying mechanisms remain largely debated [20, 61, 62, 97, 98, 108-110]. Many hypotheses have been proposed to explain the formation of the LDG [20], and these generally agree that a combination of biological processes and landscape dynamics has shaped the emergence of the LDG [20]. Hypotheses can be generally grouped into three categories [20]: (i) *time for species accumulation*, (ii) variation in *diversification rates*, and (iii) variation in *ecological limits* [Table 1 in 20].

Tropical environments can be used to exemplify these three hypothesis categories: First, the times for species accumulation propose that since tropical environments are older, they should have more time for species accumulation, without assuming further specific ecological or evolutionary mechanisms [57-60]. Second, higher temperatures in the tropics increase metabolic and mutation rates, which could lead to faster reproductive incompatibilities among populations and higher speciation rates compared with colder environments [61, 62]. Third, the tropics are generally more productive than colder environments and greater resource availability can sustain higher abundances, and therefore a larger number of species can coexist there [63, 64, 111, 112].

We implemented one model for each of these hypotheses and simulated the spread, speciation, dispersal and extinction of terrestrial organisms over the Cenozoic. We evaluated whether the emerging patterns from these simulated mechanisms correspond to the empirical LDG, phylogenetic tree imbalance and range size frequencies computed from data of major tetrapod groups, including mammals, birds, amphibians and reptiles (Figure 3).

Figure 3. Schematic representation of the case study showing the model design with two landscapes (i.e. L1 and L2) and configurations of three models (i.e. M0, M1 and M2) (Table S1), and model evaluation and test, based on multiple patterns including: LDG, range size distributions and phylogenetic balance. Selection criteria were based on empirical data from major tetrapod groups, i.e. mammals, birds, amphibians and reptiles (Table 3).

Input landscapes

The Cenozoic (i.e. 65 Ma until the present) is considered key for the diversification of the current biota [113] and is the period during which the modern LDG is expected to have been formed [114]. In the Cenozoic, the continents assumed their modern geographic configuration [24]. Climatically, this period was characterized by a general cooling, especially in the Miocene, and ended with the climatic oscillations of the Quaternary [115]. We compiled two global paleoenvironmental landscapes (i.e. L1 and L2) for the Cenozoic at 1° and ~170 kyr of spatial and temporal resolution, respectively (Note S1, Animations S1 and S2). To account for uncertainties on paleo-reconstructions on the emerging large-scale biodiversity patterns, we

used two paleo-elevation reconstructions [116, 117] associated with two approaches to estimate the paleo-temperature of sites (Note S1). L1 had temperatures defined by Köppen bands based on the geographic distribution of lithologic indicators of climate [54]. L2 had temperature defined by a composite of benthic foraminifer isotope records over time [118] and along latitude for specific time periods [119-125]. An aridity index ranging from zero to one was computed based on the subtropical arid Köppen zone for both landscapes [54]. For details see Note S1.

Hypothesis implementation

We implemented three hypotheses explaining the emergence of the LDG as different gen3sis models. The models (i.e. M0, M1 and M2) had distinct speciation and ecological processes (Figure 3, Note S1, Table S1). All simulations were initiated with one single ancestor species spread over the entire terrestrial surface of the Earth at 65 Ma, where the temperature optimum of each population matched local site conditions. Since we focused on terrestrial organisms, aquatic sites were considered inhabitable and twice as difficult to cross as terrestrial sites. This approximates the different dispersal limitation imposed by aquatic and terrestrial sites. The spherical shape of the Earth was accounted for in distance calculations by using haversine geodesic distances. Species disperse following a Weibull distribution with shape 2 or 5 and a scale of 550, 650, 750 or 850, resulting in most values being around 500–1500 km, with rare large dispersal events above 2000 km. The *evolution* function defines the temperature niche optimum to evolve following Brownian motion. Temperature niche optima are homogenized per geographic cluster by an abundance-weighted mean after ecological processes happen. We explored three rates of niche evolution, with a standard deviation equivalent to $\pm 0.1^{\circ}\text{C}$, $\pm 0.5^{\circ}\text{C}$ and $\pm 1^{\circ}\text{C}$.

gen3sis: engine for eco-evolutionary biodiversity modelling

M0. In the implementation of the *time for species accumulation*, the *ecology* function defines the species population abundance, where the abundance increases proportionally to the distance between the population temperature niche optimum and the site temperature (Note S1). Clusters of populations that accumulated differentiation over $\tau = 12, 24, 36, 48$ and 60 will speciate, corresponding to events occurring after $2, 4, 6, 8$ and 10 myr of isolation, respectively. The divergence rate between isolated clusters was kept constant (i.e. $+1$ for every 170 kyr of isolation). Model M0, assuming *time for species accumulation*, acted as a baseline model. This means that all mechanisms present in this model were the same for M1 and M2 if not specified otherwise.

M1. In the implementation of the *diversification rates*, the speciation function applies a temperature-dependent divergence between population clusters [61, 62]. Species in warmer environments accumulate divergence between disconnected clusters of populations at a higher rate (Note S1). The rate of differentiation increase was the average site temperature of the species clusters to the power of $2, 4$ or 6 plus a constant. This created a differentiation increase of $+1.5$ for isolated clusters of a species at the warmest range and $+0.5$ at the coldest range for every 170 kyr of isolation (Note S1, Figure S1). Using $\tau = 12, 24, 36, 48$ and 60 , this corresponds to a speciation event after $1.3, 2.7, 4.0, 5.3, 6.7$ myr and after $4, 8, 12, 16, 20$ myr for the warmest and coldest species, respectively.

M2. In the implementation of the *ecological limits*, the *ecology* function includes a carrying capacity k of each site that scales with area energy and aridity in the ecology function [112, 126]. The theory of carrying capacity proposes that energy limits abundances and therefore determines how many of each species can coexist in a given place [21, 112]. If the sum of all species abundances in a site is above k , species abundances are randomly reduced across species until k is reached. We explored low and high k values using a k power-law scaling of 2 and 3 .

418 *Exploration of model parameters*

419 For each model (i.e. M0, M1 and M2) in combination with each landscape (i.e. L1 and L2), we
 420 explored a range of conservative model parameters in an interactive modelling cycle (Figure
 421 3). In addition, we explored dispersal distributions and parameters ranging in realized mean
 422 and 95% quantiles between less than a single cell, i.e. ~50 km for a landscape at 4°, and more
 423 than the Earth's diameter, i.e. ~12'742 km (Figure S2). Trait evolution frequency and intensity
 424 ranged from zero to one. We ran a full factorial exploration of these parameter ranges at a
 425 coarse resolution of 4° (i.e. M0 n=480, M1 n=720, M2 n=480) and compared these to empirical
 426 data. Simulations considered further: (i) had at least one speciation event; (ii) did not have all
 427 species becoming extinct; (iii) had fewer than 50'000 species; or (iv) had fewer than 10'000
 428 species cohabiting the same site at any point in time (Note S1). After parameter range
 429 exploration, we identified realistic parameters and ran a subset at 1° for high-resolution outputs
 430 (Figure 4).

431 *Correspondence with empirical data*

432 In order to explore the parameters of all three models and compare their ability to produce the
 433 observed biodiversity patterns, we used a pattern-oriented modelling (POM) approach [23,
 434 86]. POM compares the predictions of each model and parameter combination with a number
 435 of diagnostic patterns from empirical observations. In our case, we used the LDG slope, tree
 436 imbalance and range size frequencies as diagnostics patterns (Figure 3, Note S1). The POM
 437 approach allows a calibration and model comparison based on high-level diagnostic patterns,
 438 avoiding the hurdles of defining explicit (approximate) likelihood functions [127]. The POM
 439 approach requires the specification of a range for each pattern under which observation and
 440 prediction are accepted, hence when a simulation satisfactorily reproduces empirical
 441 observations. Unless POM is coupled with an explicit probabilistic model [127], the limits for

acceptance must be decided by the modeller based on their understanding of the data [23, 86].

To generate the empirical values for these patterns, we obtained distribution data on 25'941 species [128-130], following [131], and phylogenetic data on 18'978 species [5, 132-135], following [136] for major tetrapod groups, i.e. terrestrial mammals, birds, amphibians and reptiles (Note S1). LDG is given by the percentage of species loss per latitudinal degree and measured by the slope of a linear regression on normalized species richness against absolute latitude. β -statistics [31] was used for phylogenetic tree imbalance in ultrametric trees, following [102]. Species ranges decrease (SRD) in km² is given by the percentage of species loss per species range and is measured by the slope of a linear regression of range size distributions. Empirical values of LDG, β and SRD were: mammals (LDG=5.1%, β =-0.4, SRD=2.3*10³%), birds (LDG=1.5%, β =-1.3, SRD=6.5*10⁷%), amphibians (LDG=3.9%, β =-0.7, SRD=0.11%) and reptiles (LDG=1.5%, β =-0.8, SRD=5.3*10³%). Based on these values, we used the following acceptance criteria: (i) LDG between 5.4% and 1.1%, (ii) tree shape statistic β between -1.4 and -0.3, and (iii) range size frequencies with a decrease in the number of large-range species with a tolerance of 5% [93-95] (Note S1).

Simulations results and synthesis

We found that model M2 was the best match for all the empirical patterns individually, and the only model able to pass all acceptance criteria (Table 3). Although all three models were able to reproduce the LDG, M2 was superior in explaining the LDG, phylogenetic tree imbalance and species range size frequencies simultaneously (Table 3). Most simulations of model M2 (67%) resulted in a decrease in species richness at higher latitudes, indicating that the LDG emerged systematically under M2 mechanisms (Figure S3, Tables S2, S3 and S4). Increasing the spatial resolution of the simulations (n=12) resulted in an increase in γ richness and computation time and a slight decrease of the LDG (Figure S5), which was associated with a

disproportionally larger number of sites towards higher latitudes, which also affects population connectivity and therefore speciation rates [137]. We then selected the best matching simulation of M2 in L1 at 1° (n=12) that predicted realistic biodiversity patterns (Figure 4, Animation S4). The emerging LDG (i.e. 4.6% of species loss per latitudinal degree) closely matched empirical curves, with good agreement for mammals (Pearson $r=0.6$), birds ($r=0.57$), amphibians ($r=0.57$) and reptiles ($r=0.38$) (Note S1, Figure 4C, Figure S6). Finally, we found that the support for M2 over M0 and M1 was consistent across the two alternative landscapes L1 and L2 (Figure S3, Table S4).

Our sensitivity analyses of parameters further provided information about the role of dispersal and ecological processes in shaping the LDG (Note S1, Table S2 and S3). In particular, our results indicate that an increase in the scaling factor of carrying capacity with energy k resulted in a steeper LDG slope, which is in agreement with findings from previous studies [21, 61, 112, 126]. Similarly, increasing the time for divergence consistently led to lower species richness and flattened the LDG slope so that the tropics accumulated diversity more slowly, but changes in speciation rates were less likely to drive large-scale biodiversity patterns [110]. Saupe and colleagues [25] showed that simulations with poor dispersal are better at representing the observed strong LDG in tetrapods. In agreement with their results, our parameter explorations indicated that dispersal correlated negatively with LDG [25], and simulations with lower dispersal parameters agreed better with the data (Note S1). While previous case studies [25, 26, 44] have been carried out to investigate the formation of the LDG using computer models, they used a shorter timeframe (i.e. below 1 Ma) and/or explored few mechanisms, i.e. simplified landscape or single acceptance criteria [24, 41, 42, 110]. Beyond this illustrative case study, future analyses could combine multiple mechanisms in relation to additional biodiversity patterns in order to investigate the most likely combination of mechanisms shaping the intriguing LDG pattern.

gen3sis: engine for eco-evolutionary biodiversity modelling

A Landscape (L1)

B Emerging patterns (Simulation M2 in L1)

C Correspondence with empirical data (LDG)

Figure 4. Illustration of one global simulation of the speciation, dispersal and extinction of lineages over the Cenozoic, starting with a single ancestor species and imposed energetic carrying capacity (M2 in L1). (A) Images of the Earth land masses through time, used as input for the simulation. (B) Selected emerging patterns: evolutionary dynamics; phylogeny; and present richness. (B.1) Evolutionary dynamics: γ richness ($\log_{10}$ scale) through time (blue line)

gen3sis: engine for eco-evolutionary biodiversity modelling

and diversification rate. (B.2) Phylogeny showing the distribution of the temperature optima for all extant species. (B.3) Present distribution of simulated α biodiversity globally, which indicates locations of biodiversity hotspots. For the empirical match see Figure S3. (C) Model correspondence with empirical data of terrestrial mammals, birds, amphibians and reptiles for the LDG, measured as the standardized and area-scaled mean species number per latitudinal degree.

Table 3. Model acceptance table with pattern descriptions and details of acceptance derived from empirical data. Percentages of accepted simulations (for both landscapes) are shown for each model and acceptance parameter and the combination of all acceptance patterns.

Acceptance		M0	M1	M2
Pattern	Description and empirical acceptance	(n=480)	(n=720)	(n=480)
LDG	Percentage of species loss per latitudinal degree from linear regression slope. Accept LDGs between 5% and 1%	34%	36%	42%
Phytogenic balance	The imbalance of a phylogenetic tree is measured by the value that maximizes the likelihood in the β -splitting model [138]. Accept phylogenies with β between -1.4 and -0.3	58%	51%	66%
Range	Range size distributions. Accept only distributions that show a consistent frequency decrease towards large-ranged species with a tolerance of 5%	0%	0%	5%
Combined	Simulations passing all criteria above with at least 100 species alive at the present.	0%	0%	1%

Discussion

Understanding the emergence of biodiversity patterns requires the consideration of multiple biological processes and abiotic forces that potentially underpin them [20, 26, 35, 36]. We have introduced gen3sis, a modular, spatially-explicit, eco-evolutionary simulation engine implemented as an R-package, which offers the possibility to explore ecological and macroevolutionary dynamics over changing landscapes. Gen3sis generates commonly observed diversity patterns and, thanks to its flexibility, enables the testing of a broad range of hypotheses (Table 4). It follows the principle of computer models from other fields [139-141], where mechanisms are implemented in a controlled numeric environment and emerging patterns can be compared with empirical data [23]. The combination of exploring patterns emerging from models and matching qualitatively and quantitatively the model outputs to empirical data should increase our understanding of the processes underlying global biodiversity patterns.

Using a case study, we have illustrated the flexibility and utility of gen3sis in modelling multiple eco-evolutionary hypotheses in global paleo-environmental reconstructions (Figures 3 and 4). Our findings suggest that global biodiversity patterns can be modelled realistically by combining paleo-environmental reconstructions with eco-evolutionary processes, thus moving beyond pattern description to pattern reproduction [35]. Nevertheless, in our case study we only implemented a few of the standing LDG hypotheses [20, 34]. Multiple macroecological and macroevolutionary hypotheses still have to be tested, including the role of stronger biotic interactions in the tropics than in other regions [142], and compared with more biodiversity patterns [20]. Considering multiple additional biodiversity patterns will allow a more robust selection of models. Apart from the global LDG case study, we propose an additional case study (Note S2, Figure S7) illustrating how gen3sis can be used for regional and theoretical studies, such as investigations of the effect of island ontology on the temporal dynamics of biodiversity [39, 143]. Further, illustrations associated with the programming code are offered

gen3sis: engine for eco-evolutionary biodiversity modelling

as a vignette of the R-package, which will support broad application of gen3sis. Altogether, our examples illustrate the great potential for exploration provided by gen3sis, promising future advances in our understanding of empirical biodiversity patterns.

Verbal explanations of the main principles underlying the emergence of biodiversity are frequently proposed but are rarely quantified or readily generalized across study systems [20]. We anticipate that gen3sis will be particularly useful for exploring the consequences of mechanisms that so far have mostly been verbally defined. For example, the origins of biodiversity gradients have been associated with a variety of mechanisms [7], but these represent verbal abstractions of biological processes that are difficult to evaluate [20]. Whereas simulation models can always be improved, their formulation implies formalizing process-based abstractions via mechanisms expected to shape the emergent properties of a system [144]. Specifically, when conveying models with gen3sis, decisions regarding the biological processes and landscapes must be formalized in a reproducible fashion. By introducing gen3sis, we encourage a standardization of configuration and landscape objects, which will facilitate future model comparisons. This standardization offers a robust framework for developing, testing, comparing, and applying the mechanisms relevant to biodiversity research.

Studying multiple patterns is a promising approach in disentangling competing hypotheses [20, 86]. A wide range of biodiversity dimensions can be simulated with gen3sis (Table 2), which – after appropriate sampling [145] – can serve in a multi-dimensional comparison with empirical data, i.e. a time series of species abundance matrices and trait matrices, as well as a phylogeny. These output objects are compatible with most R-packages used for community or phylogenetic analyses. Hence, the model outputs can be linked to packages computing diversification rates [146], community phylogenetics [147], or functional diversity [148]. The comparison of simulation outputs with empirical data requires a systematic exploration of processes, when formulating models, and parameter values [e.g. 149]. First, a set of mechanisms and/or a range of reasonable parameter values are explored, for instance

gen3sis: engine for eco-evolutionary biodiversity modelling

dispersal distances from measurements in a specific clade [150] and/or evolutionary rates [151]. A range of simulation outputs can then be evaluated quantitatively by studying the range of models and parameter values that produce the highest level of agreement with multiple types of empirical data, using for example a pattern-oriented modelling approach [86]. For each model, patterns are evaluated given an acceptance criteria [e.g. 40]. A multi-scale and multi-pattern comparison of simulations with empirical data can be completed to evaluate a model's ability to simultaneously reproduce not only one, but a diverse set of empirical patterns across multiple biodiversity dimensions.

The quality of the outputs of simulation models such as gen3sis hinges on accurate reconstructions of past environmental conditions [3, 42]. Although recent studies using realistic landscapes and computer models reproduced biodiversity patterns over a time scale spanning the Quaternary [25, 26, 44], many speciation and extinction events shaping present diversity patterns date back before the glaciation, and few studies have covered deep-time dynamics [24, 41, 42, 131]. Deep-time landscape reconstructions are still generally lacking but are increasingly becoming available [116, 118]. Here, we used available paleo-elevation models [116, 117] and paleoclimate indicators [54, 118-125, 152-154] to generate input landscapes to explore the formation of the LDG and account for uncertainties and limitations. For instance, we represented Quaternary climatic oscillation using ~170 kyr time-steps, which correspond to a coarser temporal scale compared with the frequency of oscillations, and thus do not account for shorter climatic variation effects on diversity patterns [25, 26, 44]. We also did not consider ice cover, that can mask species' habitable sites, which probably explains the mismatch between simulated and empirical LDG patterns below 50° (Figure 4C). Moreover, paleo indicators of climate from Köppen bands have major limitations, and the temperature estimation derived in our case study can suffer from large inaccuracies. Lastly, extrapolation of the current temperature lapse rate along elevation might lead to erroneous estimates, especially in terms of the interaction with air moisture [155], which was not further investigated here. Hence, the presented case study represents a preliminary attempt for illustrative

gen3sis: engine for eco-evolutionary biodiversity modelling

purposes. Further research is required to generate more accurate paleolandscapes, and research in biology should improve empirical evidence and our understanding of mechanisms. We expect that gen3sis will support exciting interdisciplinary research across the fields of geology, climatology and biology to understand the shaping of biodiversity.

Table 4. A non-exhaustive list of expected applications of gen3sis. Given the flexibility and the range of outputs produced by the engine, we expect that gen3sis will serve a large range of purposes, from testing a variety of theories and hypotheses to evaluating phylogenetic diversification methods.

Use	Examples from Figure 1
Testing phylogenetic inference methods, including diversification rates in phylogeographic reconstructions.	Infer diversification rate in gen3sis simulated phylogenies (E) and compare with a known diversification in gen3sis (A, B & G).
Providing biotic scenarios for past responses to geodynamics.	Based on model outputs (C–F) and comparisons with empirical data (H), select plausible models (B).
Testing paleo-climatic and paleo-topographic reconstructions using biodiversity data.	Based on model outputs (C–F) and comparisons with empirical data (H), select plausible landscape(s) (A).
Comparing expectations of different processes relating to the origin of biodiversity; generating and testing hypotheses.	Compare models (A, B & G) with outputs (C–F) and possibly how well outputs match empirical data (H).
Comparing simulated intra-specific population structure with empirical genetic data.	Compare simulated divergence matrices with population genetic data.
Forecasting the response of biodiversity to global changes (e.g. climate or fragmentation).	Extrapolate plausible and validated models (A, B & G) on landscapes under climate change scenarios (A).
Investigating trait evolution through space and time.	Combine past and present simulated species traits (F) and distributions (C, D) with fossil and trait data (H).
Modelling complex systems in space and time in unconventional biological contexts in order to investigate eco-evolutionary processes in fields traditionally not relying on biological principles.	Model eco-evolutionary mechanisms (A, B & G) in an unconventional eco-evolutionary context.

Conclusions

Here we have introduced gen3sis, a modular simulation engine that enables exploration of the consequences of ecological and evolutionary processes and feedbacks on the emergence of spatio-temporal macro-eco-evolutionary biodiversity dynamics. This modelling approach bears similarity with other computer models that have led to significant progress in other fields, such as climatology [139], cosmology [140] and conservation [141]. We showcase the versatility and utility of gen3sis by comparing the ability of three alternative mechanisms in two landscapes to generate the LDG while accounting for other global biodiversity patterns. Besides the LDG, frontiers on the origins of biodiversity involve [16]: (i) quantifying speciation, extinction and dispersal events [114]; (ii) exploring adaptive niche evolution [26, 44]; and (iii) investigating multiple diversity-dependence and carrying capacity mechanisms [21, 111, 112]. Further exploration possibilities may include: (iv) revealing the mechanisms behind age-dependent speciation and extinction patterns [102, 108, 156]; (v) contrasts between terrestrial and aquatic ecosystems [16]; and (vi) calculations of uncertainty resulting from climatic and geological dynamics [e.g. 24, 25, 26, 41, 42]. Gen3sis can support these research frontiers as a general tool for formalizing and studying existing theories associated with the origin of biodiversity, testing new hypotheses against data, and making predictions about future biodiversity trajectories (Table 4). Openly available as an R-package, gen3sis has the potential to catalyse interdisciplinary biodiversity research. We call for the formation of a community of ecologists, biologists, mathematicians, geologists, climatologists and scientists from other fields around this class of eco-evolutionary simulation models in order to unravel the processes that have shaped Earth's biodiversity.

Availability

Gen3sis is implemented in a mix of R and C++ code, and wrapped into an R-package. All high-level functions that the user may interact with are written in R, and are documented via the standard R / Roxygen help files for R-packages. Runtime-critical functions are implemented in C++ and coupled to R via the Rcpp framework. Additionally, the package provides several convenience functions to generate input data, configuration files and plots, as well as tutorials in the form of vignettes that illustrate how to declare models and run simulations. The software, under an open and free GPL3 license, can be downloaded from CRAN at <https://CRAN.R-project.org/package=gen3sis>. The development version, open to issue reporting and feature suggestions, is available at <https://github.com/project-Gen3sis/R-package>. Supporting information, such as notes, scripts, data, figures and animations, are available at <https://github.com/ohagen/SupplementaryInformationGen3sis>, facilitating full reproducibility.

Acknowledgements

We thank Samuel Bickel and Alex Skeels for thorough comments on this manuscript and package. We thank Camille Albouy, Charles N.D. Santana, Lydian Boschman, Wilhelmine Bach, Thomas Keggin, Flurin Leugger, Victor L.J. Boussange, Conor Waldock and all sELDiG working group participants for insightful feedback during the model development. We thank the WSL and ETH Zürich for support and infrastructure including access to High Performance Computing facilities.

640 References

- 641 1. Darwin C. On the origin of species by means of natural selection, or preservation
642 of favoured races in the struggle for life. London: John Murray; 1859.
- 643 2. Couvreur TLP, Dauby G, Blach-Overgaard A, Deblauwe V, Dessein S, Droissart V,
644 et al. Tectonics, climate and the diversification of the tropical African terrestrial flora
645 and fauna. Biological Reviews Cambridge Philosophical Society. 2020. Epub
646 2020/09/15. doi: 10.1111/brv.12644.
- 647 3. Pellissier L, Heine C, Rosauer DF, Albouy C. Are global hotspots of endemic
648 richness shaped by plate tectonics? Biol J Linn Soc. 2017;123(1):247-61. doi:
649 10.1093/biolinnean/blx125.
- 650 4. Gaston KJ. Global patterns in biodiversity. Nature. 2000;405(6783):220-7. doi:
651 10.1038/35012228.
- 652 5. Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO. The global diversity of birds
653 in space and time. Nature. 2012;491(7424):444-8. Epub 2012/11/06. doi:
654 10.1038/nature11631.
- 655 6. Mittelbach GG, Schemske DW, Cornell HV, Allen AP, Brown JM, Bush MB, et al.
656 Evolution and the latitudinal diversity gradient: speciation, extinction and
657 biogeography. Ecol Lett. 2007;10(4):315-31. doi: 10.1111/j.1461-0248.2007.01020.x.
- 658 7. Fine PVA. Ecological and evolutionary drivers of geographic variation in species
659 diversity. Annual Review of Ecology, Evolution, and Systematics. 2015;46(1):369-92.
660 doi: 10.1146/annurev-ecolsys-112414-054102.
- 661 8. Humboldt Av, Bonpland A. Ideen zu einer geographie der pflanzen nebst einem
662 naturgemälde der tropenländer, auf beobachtungen und messungen gegründet.
663 Tübingen.; F. G. Cotta; etc.; 1807. 2 p.l., xii, 182 p. p.
- 664 9. Wallace AR. Tropical nature, and other essays. London: Macmillan and Company;
665 1878.
- 666 10. Hutchinson GE. Homage to Santa Rosalia or why are there so many kinds of
667 animals? The American Naturalist. 1959;93(870):145-59. doi: 10.1086/282070.
- 668 11. MacArthur R, Wilson E. The theory of island biogeography. Princeton, NJ:
669 Princeton Univ. Press; 1967.
- 670 12. Pianka ER. Convexity, desert lizards, and spatial heterogeneity. Ecology.
671 1966;47(6):1055-9. doi: 10.2307/1935656.
- 672 13. Willig MR, Kaufman DM, Stevens RD. Latitudinal gradients of biodiversity:
673 Pattern, process, scale, and synthesis. Annual Review of Ecology, Evolution, and
674 Systematics. 2003;34(1):273-309. doi: 10.1146/annurev.ecolsys.34.012103.144032.
- 675 14. Schemske DW, Mittelbach GG. "Latitudinal gradients in species diversity":
676 Reflections on Pianka's 1966 article and a look forward. The American Naturalist.
677 2017;189(6):599-603. Epub 2017/05/18. doi: 10.1086/691719.
- 678 15. MacArthur RH. Patterns of species diversity. Biological Reviews. 1965;40(4):510-
679 33. doi: 10.1111/j.1469-185X.1965.tb00815.x.
- 680 16. Benton MJ. Origins of Biodiversity. PLoS Biol. 2016;14(11):e2000724. doi:
681 10.1371/journal.pbio.2000724.
- 682 17. Morlon H, Parsons TL, Plotkin JB. Reconciling molecular phylogenies with the
683 fossil record. Proceedings of the National Academy of Sciences. 2011;108(39):16327-
684 32. Epub 2011/09/21. doi: 10.1073/pnas.1102543108.

gen3sis: engine for eco-evolutionary biodiversity modelling

18. Field R, Hawkins BA, Cornell HV, Currie DJ, Diniz-Filho JAF, Guégan J-F, et al. Spatial species-richness gradients across scales: a meta-analysis. *J Biogeogr.* 2009;36(1):132-47. doi: 10.1111/j.1365-2699.2008.01963.x.
19. Sandel B, Arge L, Dalsgaard B, Davies RG, Gaston KJ, Sutherland WJ, et al. The influence of Late Quaternary climate-change velocity on species endemism. *Science.* 2011;334(6056):660-4.
20. Pontarp M, Bunnefeld L, Cabral JS, Etienne RS, Fritz SA, Gillespie R, et al. The latitudinal diversity gradient: Novel understanding through mechanistic eco-evolutionary models. *Trends in Ecology and Evolution.* 2019;34(3):211-23. doi: 10.1016/j.tree.2018.11.009.
21. Etienne RS, Cabral JS, Hagen O, Hartig F, Hurlbert AH, Pellissier L, et al. A minimal model for the latitudinal diversity gradient suggests a dominant role for ecological limits. *The American Naturalist.* 2019;194(5):E122-E33. Epub 2019/10/16. doi: 10.1086/705243.
22. Doebeli M, Dieckmann U. Speciation along environmental gradients. *Nature.* 2003;421(6920):259-64. doi: 10.1038/nature01274.
23. Grimm V, Revilla E, Berger U, Jeltsch F, Mooij WM, Railsback SF, et al. Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. *Science.* 2005;310(5750):987-91. doi: 10.1126/science.1116681.
24. Leprieur F, Descombes P, Gaboriau T, Cowman PF, Parravicini V, Kulbicki M, et al. Plate tectonics drive tropical reef biodiversity dynamics. *Nature Communications.* 2016;7:11461. doi: 10.1038/ncomms11461.
25. Saupe EE, Myers CE, Townsend Peterson A, Soberon J, Singarayer J, Valdes P, et al. Spatio-temporal climate change contributes to latitudinal diversity gradients. *Nature Ecology & Evolution.* 2019;3(10):1419-29. Epub 2019/09/11. doi: 10.1038/s41559-019-0962-7.
26. Rangel TF, Edwards NR, Holden PB, Diniz-Filho JAF, Gosling WD, Coelho MTP, et al. Modeling the ecology and evolution of biodiversity: Biogeographical cradles, museums, and graves. *Science.* 2018;361(6399):eaar5452. doi: 10.1126/science.aar5452.
27. Duputie A, Massol F. An empiricist's guide to theoretical predictions on the evolution of dispersal. *Interface Focus.* 2013;3(6):20130028. Epub 2014/02/12. doi: 10.1098/rsfs.2013.0028.
28. Keller I, Seehausen O. Thermal adaptation and ecological speciation. *Mol Ecol.* 2012;21(4):782-99. Epub 2011/12/21. doi: 10.1111/j.1365-294X.2011.05397.x.
29. Urban MC. The evolution of species interactions across natural landscapes. *Ecol Lett.* 2011;14(7):723-32. Epub 2011/05/28. doi: 10.1111/j.1461-0248.2011.01632.x.
30. Mittelbach GG, Schemske DW. Ecological and evolutionary perspectives on community assembly. *Trends in Ecology and Evolution.* 2015;30(5):241-7. Epub 2015/03/26. doi: 10.1016/j.tree.2015.02.008.
31. Schluter D, Pennell MW. Speciation gradients and the distribution of biodiversity. *Nature.* 2017;546(7656):48-55. doi: 10.1038/nature22897.
32. Villa Martin P, Hidalgo J, Rubio de Casas R, Munoz MA. Eco-evolutionary Model of Rapid Phenotypic Diversification in Species-Rich Communities. *PLoS Comp Biol.* 2016;12(10):e1005139. doi: 10.1371/journal.pcbi.1005139.
33. Leidinger L, Cabral J. Biodiversity dynamics on islands: Explicitly accounting for causality in mechanistic models. *Diversity.* 2017;9(3). doi: 10.3390/d9030030.

34. Pontarp M, Wiens JJ. The origin of species richness patterns along environmental gradients: uniting explanations based on time, diversification rate and carrying capacity. *J Biogeogr.* 2017;44(4):722-35. doi: 10.1111/jbi.12896.
35. Gotelli NJ, Anderson MJ, Arita HT, Chao A, Colwell RK, Connolly SR, et al. Patterns and causes of species richness: a general simulation model for macroecology. *Ecol Lett.* 2009;12(9):873-86. doi: 10.1111/j.1461-0248.2009.01353.x.
36. Cabral JS, Valente L, Hartig F. Mechanistic simulation models in macroecology and biogeography: state-of-art and prospects. *Ecography.* 2017;40(2):267-80. doi: 10.1111/ecog.02480.
37. Connolly SR, Keith SA, Colwell RK, Rahbek C. Process, mechanism, and modeling in macroecology. *Trends in Ecology and Evolution.* 2017;32(11):835-44. doi: 10.1016/j.tree.2017.08.011.
38. Skeels A, Cardillo M. Reconstructing the geography of speciation from contemporary biodiversity data. *The American Naturalist.* 2019;193(2):240-55. Epub 2019/02/06. doi: 10.1086/701125.
39. Cabral JS, Wiegand K, Kreft H. Interactions between ecological, evolutionary and environmental processes unveil complex dynamics of insular plant diversity. *J Biogeogr.* 2019. doi: 10.1111/jbi.13606.
40. Cabral JS, Whittaker RJ, Wiegand K, Kreft H, Emerson B. Assessing predicted isolation effects from the general dynamic model of island biogeography with an eco-evolutionary model for plants. *J Biogeogr.* 2019. doi: 10.1111/jbi.13603.
41. Donati GFA, Parravicini V, Leprieur F, Hagen O, Gaboriau T, Heine C, et al. A process-based model supports an association between dispersal and the prevalence of species traits in tropical reef fish assemblages. *Ecography.* 2019;42(12):2095-106. doi: 10.1111/ecog.04537.
42. Saupe EE, Qiao H, Donnadieu Y, Farnsworth A, Kennedy-Asser AT, Ladant J-B, et al. Extinction intensity during Ordovician and Cenozoic glaciations explained by cooling and palaeogeography. *Nature Geoscience.* 2020;13(1):65-70. doi: 10.1038/s41561-019-0504-6.
43. Gaboriau T, Albouy C, Descombes P, Mouillot D, Pellissier L, Leprieur F. Ecological constraints coupled with deep-time habitat dynamics predict the latitudinal diversity gradient in reef fishes. *Proceedings of the Royal Society B: Biological Sciences.* 2019;286(1911):20191506. Epub 2019/09/19. doi: 10.1098/rspb.2019.1506.
44. Saupe EE, Myers CE, Peterson AT, Soberón J, Singarayer J, Valdes P, et al. Non-random latitudinal gradients in range size and niche breadth predicted by spatial patterns of climate. *Global Ecol Biogeogr.* 2019;28(7):928-42. doi: 10.1111/geb.12904.
45. Pontarp M, Brännström Å, Petchey OL, Poisot T. Inferring community assembly processes from macroscopic patterns using dynamic eco-evolutionary models and Approximate Bayesian Computation (ABC). *Methods in Ecology and Evolution.* 2019;10(4):450-60. doi: 10.1111/2041-210x.13129.
46. Sukumaran J, Economo EP, Lacey Knowles L. Machine Learning Biogeographic Processes from Biotic Patterns: A New Trait-Dependent Dispersal and Diversification Model with Model Choice By Simulation-Trained Discriminant Analysis. *Syst Biol.* 2016;65(3):525-45. Epub 2015/12/31. doi: 10.1093/sysbio/syv121.
47. Kubisch A, Holt RD, Poethke H-J, Fronhofer EA. Where am I and why? Synthesizing range biology and the eco-evolutionary dynamics of dispersal. *Oikos.* 2014;123(1):5-22. doi: 10.1111/j.1600-0706.2013.00706.x.

- 779 48. Pigot AL, Phillimore AB, Owens IP, Orme CD. The shape and temporal dynamics
780 of phylogenetic trees arising from geographic speciation. *Syst Biol.* 2010;59(6):660-73.
781 Epub 2010/10/19. doi: 10.1093/sysbio/syq058.
- 782 49. Rangel Thiago Fernando LVB, Diniz-Filho José Alexandre F, Colwell Robert K.
783 Species Richness and Evolutionary Niche Dynamics: A Spatial Pattern-Oriented
784 Simulation Experiment. *The American Naturalist.* 2007;170(4):602-16. doi:
785 10.1086/521315.
- 786 50. Rangel TFLVB, Diniz-Filho JAF. An evolutionary tolerance model explaining
787 spatial patterns in species richness under environmental gradients and geometric
788 constraints. *Ecography.* 2005;28(2):253-63. doi: 10.1111/j.0906-7590.2005.04038.x.
- 789 51. Culina A, van den Berg I, Evans S, Sanchez-Tojar A. Low availability of code in
790 ecology: A call for urgent action. *PLoS Biol.* 2020;18(7):e3000763. Epub 2020/07/30.
791 doi: 10.1371/journal.pbio.3000763.
- 792 52. Muneeppeerakul R, Bertuzzo E, Rinaldo A, Rodriguez-Iturbe I. Evolving
793 biodiversity patterns in changing river networks. *J Theor Biol.* 2019;462:418-24. Epub
794 2018/11/30. doi: 10.1016/j.jtbi.2018.11.021.
- 795 53. Jöks M, Pärtel M. Plant diversity in Oceanic archipelagos: realistic patterns
796 emulated by an agent-based computer simulation. *Ecography.* 2018;42(4):740-54. doi:
797 10.1111/ecog.03985.
- 798 54. Hagen O, Vaterlaus L, Albouy C, Brown A, Leugger F, Onstein RE, et al. Mountain
799 building, climate cooling and the richness of cold-adapted plants in the Northern
800 Hemisphere. *J Biogeogr.* 2019. doi: 10.1111/jbi.13653.
- 801 55. Xu X, Kuntner M, Liu F, Chen J, Li D. Formation of rivers and mountains drives
802 diversification of primitively segmented spiders in continental East Asia. *J Biogeogr.*
803 2018. doi: 10.1111/jbi.13403.
- 804 56. Ficetola GF, Mazel F, Thuiller W. Global determinants of zoogeographical
805 boundaries. *Nature Ecology & Evolution.* 2017;1:0089. doi: 10.1038/s41559-017-0089.
- 806 57. Wiens JJ, Donoghue MJ. Historical biogeography, ecology and species richness.
807 *Trends in Ecology and Evolution.* 2004;19(12):639-44. doi: 10.1016/j.tree.2004.09.011.
- 808 58. Wiens JJ, Graham CH. Niche conservatism: Integrating evolution, ecology, and
809 conservation biology. *Annual Review of Ecology, Evolution, and Systematics.*
810 2005;36(1):519-39. doi: 10.1146/annurev.ecolsys.36.102803.095431.
- 811 59. Fischer AG. Latitudinal Variations in Organic Diversity. *Evolution.*
812 1960;14(1):64-81. doi: 10.2307/2405923.
- 813 60. Patrick R. Stephens, John J. Wiens. Explaining Species Richness from Continents
814 to Communities: The Time-for-Speciation Effect in Emydid Turtles. *The American*
815 *Naturalist.* 2003;161(1):112-28. doi: 10.1086/345091.
- 816 61. Allen AP, Gillooly JF, Savage VM, Brown JH. Kinetic effects of temperature on
817 rates of genetic divergence and speciation. *Proceedings of the National Academy of*
818 *Sciences.* 2006;103(24):9130-5. doi: 10.1073/pnas.0603587103.
- 819 62. Rohde K. Latitudinal gradients in species diversity: the search for the primary
820 cause. *Oikos.* 1992;514-27.
- 821 63. MacArthur RH, Wilson EO. An Equilibrium Theory of Insular Zoogeography.
822 *Evolution.* 1963;17(4):373-87. doi: 10.2307/2407089.
- 823 64. Rosenzweig ML. Species diversity in space and time: Cambridge University Press;
824 1995.
- 825 65. R Core Team. R: A Language and Environment for Statistical Computing. In:
826 Computing RFFS, editor. Vienna, Austria2020.

66. Bivand RS. Progress in the R ecosystem for representing and handling spatial data. *Journal of Geographical Systems*. 2020. doi: 10.1007/s10109-020-00336-0.

67. Wilson G, Aruliah DA, Brown CT, Chue Hong NP, Davis M, Guy RT, et al. Best practices for scientific computing. *PLoS Biol*. 2014;12(1):e1001745. doi: 10.1371/journal.pbio.1001745.

68. Alzate A, Janzen T, Bonte D, Rosindell J, Etienne RS, Belmaker J. A simple spatially explicit neutral model explains the range size distribution of reef fishes. *Global Ecol Biogeogr*. 2019;28(7):875-90. doi: 10.1111/geb.12899.

69. Salles T, Rey P, Bertuzzo E. Mapping landscape connectivity as a driver of species richness under tectonic and climatic forcing. *Earth Surface Dynamics*. 2019;7(4):895-910. doi: 10.5194/esurf-7-895-2019.

70. Giezendanner J, Bertuzzo E, Pasetto D, Guisan A, Rinaldo A. A minimalist model of extinction and range dynamics of virtual mountain species driven by warming temperatures. *PLoS One*. 2019;14(3):e0213775. Epub 2019/03/19. doi: 10.1371/journal.pone.0213775.

71. van Etten J. R package gdistance: Distances and routes on geographical grids. *Journal of Statistical Software*. 2017;76(13). doi: 10.18637/jss.v076.i13.

72. Dobzhansky T. *Genetics and the origin of species*. Repr. ed. New York: Columbia University Press; 1982. 364 S. p.

73. Stanley SM. Population Size, Extinction, and Speciation: The Fission Effect in Neogene Bivalvia. *Paleobiology*. 1986;12(1):89-110.

74. Wisz MS, Pottier J, Kissling WD, Pellissier L, Lenoir J, Damgaard CF, et al. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews Cambridge Philosophical Society*. 2013;88(1):15-30. doi: 10.1111/j.1469-185X.2012.00235.x.

75. Pontarp M, Petchey OL. Ecological opportunity and predator-prey interactions: linking eco-evolutionary processes and diversification in adaptive radiations. *Proceedings of the Royal Society B: Biological Sciences*. 2018;285(1874). Epub 2018/03/09. doi: 10.1098/rspb.2017.2550.

76. Pellissier L, Descombes P, Hagen O, Chalmandrier L, Glauser G, Kergunteuil A, et al. Growth-competition-herbivore resistance trade-offs and the responses of alpine plant communities to climate change. *Funct Ecol*. 2018;32(7):1693-703. doi: 10.1111/1365-2435.13075.

77. Pellissier L. Stability and the competition-dispersal trade-off as drivers of speciation and biodiversity gradients. *Frontiers in Ecology and Evolution*. 2015;3. doi: 10.3389/fevo.2015.00052.

78. Ronce O. How Does It Feel to Be Like a Rolling Stone? Ten Questions About Dispersal Evolution. *Annual Review of Ecology, Evolution, and Systematics*. 2007;38(1):231-53. doi: 10.1146/annurev.ecolsys.38.091206.095611.

79. Stegen JC, Enquist BJ, Ferriere R. Advancing the metabolic theory of biodiversity. *Ecol Lett*. 2009;12(10):1001-15. doi: 10.1111/j.1461-0248.2009.01358.x.

80. Gillooly JF, Allen AP, West GB, Brown JH. The rate of DNA evolution: Effects of body size and temperature on the molecular clock. *Proceedings of the National Academy of Sciences of the United States of America*. 2005;102(1):140-5. doi: 10.1073/pnas.0407735101.

81. Ohta T. The nearly neutral theory of molecular evolution. *Annu Rev Ecol Syst*. 1992;23(1):263-86. doi: pdf/10.1146/annurev.es.23.110192.001403.

82. Felsenstein J. Maximum-likelihood estimation of evolutionary trees from continuous characters. *Am J Hum Genet*. 1973;25(5):471.

- 876 83. Hansen TF. Stabilizing selection and the comparative analysis of adaptation.
877 Evolution. 1997;51(5):1341-51.
- 878 84. Macarthur R, Levins R. The limiting similarity, convergence, and divergence of
879 coexisting species. The American Naturalist. 1967;101(921):377-85. doi:
880 10.1086/282505.
- 881 85. Cabral JS, Kreft H, Higgins S. Linking ecological niche, community ecology and
882 biogeography: insights from a mechanistic niche model. J Biogeogr. 2012;39(12):2212-
883 24. doi: 10.1111/jbi.12010.
- 884 86. Grimm V, Railsback SF. Pattern-oriented modelling: a 'multi-scope' for predictive
885 systems ecology. Philosophical Transactions of the Royal Society B.
886 2012;367(1586):298-310. Epub 2011/12/07. doi: 10.1098/rstb.2011.0180.
- 887 87. Lewin R. Biologists disagree over bold signature of nature. Science.
888 1989;244(4904):527-9.
- 889 88. Lawton JH. Are there general laws in ecology? Oikos. 1999;84(2):177-92. doi:
890 10.2307/3546712.
- 891 89. Swenson NG, Anglada-Cordero P, Barone JA. Deterministic tropical tree
892 community turnover: evidence from patterns of functional beta diversity along an
893 elevational gradient. Proceedings of the Royal Society B: Biological Sciences.
894 2011;278(1707):877-84. Epub 2010/09/24. doi: 10.1098/rspb.2010.1369.
- 895 90. Anderson MJ, Crist TO, Chase JM, Vellend M, Inouye BD, Freestone AL, et al.
896 Navigating the multiple meanings of beta diversity: a roadmap for the practicing
897 ecologist. Ecol Lett. 2011;14(1):19-28. Epub 2010/11/13. doi: 10.1111/j.1461-
898 0248.2010.01552.x.
- 899 91. Holt BG, Lessard J-P, Borregaard MK, Fritz SA, Araújo MB, Dimitrov D, et al. An
900 update of wallace's zoogeographic regions of the world. Science. 2013;339(6115):74-8.
901 doi: 10.1126/science.1228282.
- 902 92. Qian H, Ricklefs RE, White PS. Beta diversity of angiosperms in temperate floras
903 of eastern Asia and eastern North America. Ecol Lett. 2005;8(1):15-22. doi:
904 10.1111/j.1461-0248.2004.00682.x.
- 905 93. Bar-On YM, Phillips R, Milo R. The biomass distribution on Earth. Proceedings of
906 the National Academy of Sciences. 2018;115(25):6506-11. Epub 2018/05/23. doi:
907 10.1073/pnas.1711842115.
- 908 94. Brown JH. On the relationship between abundance and distribution of species.
909 The American Naturalist. 1984;124(2):255-79. doi: 10.1086/284267.
- 910 95. Gaston KJ. Species-range-size distributions: patterns, mechanisms and
911 implications. Trends Ecol Evol. 1996;11(5):197-201. doi: 10.1016/0169-
912 5347(96)10027-6.
- 913 96. Gómez-Rodríguez C, Baselga A, Wiens JJ. Is diversification rate related to climatic
914 niche width? Global Ecol Biogeogr. 2015;24(4):383-95. doi: 10.1111/geb.12229.
- 915 97. Lamanna C, Blonder B, Violle C, Kraft NJ, Sandel B, Simova I, et al. Functional trait
916 space and the latitudinal diversity gradient. Proceedings of the National Academy of
917 Sciences. 2014;111(38):13745-50. Epub 2014/09/17. doi: 10.1073/pnas.1317722111.
- 918 98. Rolland J, Salamin N. Niche width impacts vertebrate diversification. Global Ecol
919 Biogeogr. 2016;25(10):1252-63. doi: 10.1111/geb.12482.
- 920 99. Morlon H. Phylogenetic approaches for studying diversification. Ecol Lett.
921 2014;17(4):508-25. Epub 2014/02/19. doi: 10.1111/ele.12251.
- 922 100. Silvestro D, Warnock RCM, Gavryushkina A, Stadler T. Closing the gap between
923 palaeontological and neontological speciation and extinction rate estimates. Nature

924 Communications. 2018;9(1):5237. Epub 2018/12/12. doi: 10.1038/s41467-018-
925 07622-y.

926 101. Stadler T. Mammalian phylogeny reveals recent diversification rate shifts.
927 Proceedings of the National Academy of Sciences. 2011;108(15):6187-92. Epub
928 2011/03/30. doi: 10.1073/pnas.1016876108.

929 102. Hagen O, Hartmann K, Steel M, Stadler T. Age-dependent speciation can explain
930 the shape of empirical phylogenies. Syst Biol. 2015;64(3):432-40. doi:
931 10.1093/sysbio/syv001.

932 103. Manceau M, Lambert A, Morlon H. Phylogenies support out-of-equilibrium
933 models of biodiversity. Ecol Lett. 2015;18(4):347-56. Epub 2015/02/26. doi:
934 10.1111/ele.12415.

935 104. Swenson NG, Erickson DL, Mi X, Bourg NA, Forero-Montaña J, Ge X, et al.
936 Phylogenetic and functional alpha and beta diversity in temperate and tropical tree
937 communities. Ecology. 2012;93(sp8):S112-S25. doi: 10.1890/11-0402.1.

938 105. Graham CH, Fine PV. Phylogenetic beta diversity: linking ecological and
939 evolutionary processes across space in time. Ecol Lett. 2008;11(12):1265-77. Epub
940 2008/12/03. doi: 10.1111/j.1461-0248.2008.01256.x.

941 106. Morlon H, Schwikl DW, Bryant JA, Marquet PA, Rebelo AG, Tauss C, et al. Spatial
942 patterns of phylogenetic diversity. Ecol Lett. 2011;14(2):141-9. Epub 2010/12/21. doi:
943 10.1111/j.1461-0248.2010.01563.x.

944 107. Petchey OL, Gaston KJ. Functional diversity: back to basics and looking forward.
945 Ecol Lett. 2006;9(6):741-58. Epub 2006/05/19. doi: 10.1111/j.1461-
946 0248.2006.00924.x.

947 108. Silvestro D, Castiglione S, Mondanaro A, Serio C, Melchionna M, Piras P, et al. A
948 450 million years long latitudinal gradient in age-dependent extinction. Ecol Lett.
949 2020;23(3):439-46. Epub 2019/12/20. doi: 10.1111/ele.13441.

950 109. Igea J, Tanentzap AJ. Angiosperm speciation cools down in the tropics. Ecol Lett.
951 2020;23(4):692-700. Epub 2020/02/12. doi: 10.1111/ele.13476.

952 110. Tittensor DP, Worm B. A neutral-metabolic theory of latitudinal biodiversity.
953 Global Ecol Biogeogr. 2016;25(6):630-41. doi: 10.1111/geb.12451.

954 111. Storch D, Bohdalkova E, Okie J. The more-individuals hypothesis revisited: the
955 role of community abundance in species richness regulation and the productivity-
956 diversity relationship. Ecol Lett. 2018;21(6):920-37. Epub 2018/04/17. doi:
957 10.1111/ele.12941.

958 112. Hurlbert AH, Stegen JC. When should species richness be energy limited, and how
959 would we know? Ecol Lett. 2014;17(4):401-13. Epub 2014/01/08. doi:
960 10.1111/ele.12240.

961 113. Kissling WD, Eiserhardt WL, Baker WJ, Borchsenius F, Couvreur TL, Balslev H, et
962 al. Cenozoic imprints on the phylogenetic structure of palm species assemblages
963 worldwide. Proceedings of the National Academy of Sciences. 2012;109(19):7379-84.
964 Epub 2012/04/25. doi: 10.1073/pnas.1120467109.

965 114. Meseguer AS, Condamine FL. Ancient tropical extinctions at high latitudes
966 contributed to the latitudinal diversity gradient. Evolution. 2020. Epub 2020/04/05.
967 doi: 10.1111/evo.13967.

968 115. Zachos J, Pagani M, Sloan L, Thomas E, Billups K. Trends, rhythms, and
969 aberrations in global climate 65Ma to present. Science. 2001;292(5517):686-93. doi:
970 10.1126/science.1059412.

116. Straume EO, Gaina C, Medvedev S, Nisancioglu KH. Global Cenozoic
Paleobathymetry with a focus on the Northern Hemisphere Oceanic Gateways.
Gondwana Research. 2020;86:126-43. doi: 10.1016/j.gr.2020.05.011.

117. Scotese CR, Wright N. PALEOMAP Paleodigital Elevation Models (PaleoDEMS) for
the Phanerozoic. 2018. Available from: <https://www.earthbyte.org/paleodem-resource-scotese-and-wright-2018>.

118. Westerhold T, Marwan N, Drury AJ, Liebrand D, Agnini C, Anagnostou E, et al. An
astronomically dated record of Earth's climate and its predictability over the last 66
million years. Science. 2020;369(6509):1383-7. doi: 10.1126/science.aba6853.

119. Zhang L, Hay WW, Wang C, Gu X. The evolution of latitudinal temperature
gradients from the latest Cretaceous through the Present. Earth-Sci Rev. 2019;189:147-
58. doi: 10.1016/j.earscirev.2019.01.025.

120. Hollis CJ, Dunkley Jones T, Anagnostou E, Bijl PK, Cramwinckel MJ, Cui Y, et al.
The DeepMIP contribution to PMIP4: methodologies for selection, compilation and
analysis of latest Paleocene and early Eocene climate proxy data, incorporating version
0.1 of the DeepMIP database. Geoscientific Model Development. 2019;12(7):3149-206.
doi: 10.5194/gmd-12-3149-2019.

121. Hutchinson DK, de Boer AM, Coxall HK, Caballero R, Nilsson J, Baatsen M. Climate
sensitivity and meridional overturning circulation in the late Eocene using GFDL CM2.1.
Climate of the Past. 2018;14(6):789-810. doi: 10.5194/cp-14-789-2018.

122. Evans D, Sagoo N, Renema W, Cotton LJ, Muller W, Todd JA, et al. Eocene
greenhouse climate revealed by coupled clumped isotope-Mg/Ca thermometry.
Proceedings of the National Academy of Sciences. 2018;115(6):1174-9. Epub
2018/01/24. doi: 10.1073/pnas.1714744115.

123. Cramwinckel MJ, Huber M, Kocken IJ, Agnini C, Bijl PK, Bohaty SM, et al.
Synchronous tropical and polar temperature evolution in the Eocene. Nature.
2018;559(7714):382-6. Epub 2018/07/04. doi: 10.1038/s41586-018-0272-2.

124. Sijp WP, von der Heydt AS, Dijkstra HA, Flögel S, Douglas PMJ, Bijl PK. The role of
ocean gateways on cooling climate on long time scales. Global Planet Change.
2014;119:1-22. doi: 10.1016/j.gloplacha.2014.04.004.

125. Keating-Bitonti CR, Ivany LC, Affek HP, Douglas P, Samson SD. Warm, not super-
hot, temperatures in the early Eocene subtropics. Geology. 2011;39(8):771-4. doi:
10.1130/g32054.1.

126. Storch D, Okie JG, Field R. The carrying capacity for species richness. Global Ecol
Biogeogr. 2019;28(10):1519-32. doi: 10.1111/geb.12987.

127. Hartig F, Calabrese JM, Reineking B, Wiegand T, Huth A. Statistical inference for
stochastic simulation models--theory and application. Ecol Lett. 2011;14(8):816-27.
Epub 2011/06/18. doi: 10.1111/j.1461-0248.2011.01640.x.

128. Jenkins CN, Pimm SL, Joppa LN. Global patterns of terrestrial vertebrate diversity
and conservation. Proceedings of the National Academy of Sciences.
2013;110(28):E2602-10. Epub 2013/06/28. doi: 10.1073/pnas.1302251110.

129. IUCN. IUCN Red List of threatened species. 2020 [cited 2020 1 Januray].
Available from: <http://www.iucnredlist.org>

130. BirdLife International. Data zone BirdLife international 2020 [cited 2020 1
January]. Available from: <http://datazone.birdlife.org>.

131. Descombes P, Leprieur F, Albouy C, Heine C, Pellissier L. Spatial imprints of plate
tectonics on extant richness of terrestrial vertebrates. J Biogeogr. 2017;44(5):1185-97.
doi: 10.1111/jbi.12959.

gen3sis: engine for eco-evolutionary biodiversity modelling

- 1019 132. Bininda-Emonds OR, Cardillo M, Jones KE, MacPhee RD, Beck RM, Grenyer R, et
1020 al. The delayed rise of present-day mammals. *Nature*. 2007;446(7135):507-12. Epub
1021 2007/03/30. doi: 10.1038/nature05634.
- 1022 133. Kuhn TS, Mooers AØ, Thomas GH. A simple polytomy resolver for dated
1023 phylogenies. *Methods in Ecology and Evolution*. 2011;2(5):427-36. doi: 10.1111/j.2041-
1024 210X.2011.00103.x.
- 1025 134. Pyron RA, Burbrink FT. Early origin of viviparity and multiple reversions to
1026 oviparity in squamate reptiles. *Ecol Lett*. 2014;17(1):13-21. Epub 2013/08/21. doi:
1027 10.1111/ele.12168.
- 1028 135. Pyron RA, Wiens JJ. Large-scale phylogenetic analyses reveal the causes of high
1029 tropical amphibian diversity. *Proceedings of the Royal Society B: Biological Sciences*.
1030 2013;280(1770):20131622. Epub 2013/09/13. doi: 10.1098/rspb.2013.1622.
- 1031 136. Meseguer AS, Antoine PO, Fouquet A, Delsuc F, Condamine FL, McGill B. The role
1032 of the Neotropics as a source of world tetrapod biodiversity. *Global Ecol Biogeogr*.
1033 2020;29(9):1565-78. doi: 10.1111/geb.13141.
- 1034 137. Rahbek C, Graves GR. Multiscale assessment of patterns of avian species richness.
1035 *Proceedings of the National Academy of Sciences*. 2001;98(8):4534-9. doi:
1036 10.1073/pnas.071034898.
- 1037 138. Aldous DJ. Stochastic models and descriptive statistics for phylogenetic trees,
1038 from Yule to today. *Statistical Science*. 2001;16(1):23-34. doi: 10.1214/ss/998929474.
- 1039 139. Simpkins G. Progress in climate modelling. *Nature Climate Change*.
1040 2017;7(10):684-5. doi: 10.1038/nclimate3398.
- 1041 140. Vogelsberger M, Marinacci F, Torrey P, Puchwein E. Cosmological simulations of
1042 galaxy formation. *Nature Reviews Physics*. 2020;2(1):42-66. doi: 10.1038/s42254-019-
1043 0127-2.
- 1044 141. Gaucherel C, Carpentier C, Geijzendorffer IR, Noûs C, Pommereau F. Discrete-
1045 event models for conservation assessment of integrated ecosystems. *Ecological*
1046 *Informatics*. 2020. doi: 10.1016/j.ecoinf.2020.101205.
- 1047 142. Schemske DW, Mittelbach GG, Cornell HV, Sobel JM, Roy K. Is There a Latitudinal
1048 Gradient in the Importance of Biotic Interactions? *Annual Review of Ecology, Evolution,*
1049 *and Systematics*. 2009;40(1):245-69. doi: 10.1146/annurev.ecolsys.39.110707.173430.
- 1050 143. Borregaard MK, Matthews TJ, Whittaker RJ, Field R. The general dynamic model:
1051 towards a unified theory of island biogeography? *Global Ecol Biogeogr*. 2016;25(7):805-
1052 16. doi: 10.1111/geb.12348.
- 1053 144. Levins R. The strategy of model building in population biology. *Am Sci*.
1054 1966;54(4):421-31.
- 1055 145. Zurell D, Berger U, Cabral JS, Jeltsch F, Meynard CN, Münkemüller T, et al. The
1056 virtual ecologist approach: simulating data and observers. *Oikos*. 2010;119(4):622-35.
1057 doi: 10.1111/j.1600-0706.2009.18284.x.
- 1058 146. Morlon H, Lewitus E, Condamine FL, Manceau M, Clavel J, Drury J, et al. RPANDA:
1059 an R package for macroevolutionary analyses on phylogenetic trees. *Methods in Ecology*
1060 *and Evolution*. 2016;7(5):589-97. doi: 10.1111/2041-210x.12526.
- 1061 147. Tsirogiannis C, Sandel B. PhyloMeasures: a package for computing phylogenetic
1062 biodiversity measures and their statistical moments. *Ecography*. 2016;39(7):709-14.
1063 doi: 10.1111/ecog.01814.
- 1064 148. Laliberté E, Legendre P, Shipley B, Laliberté ME. Package 'FD': Measuring
1065 functional diversity from multiple traits, and other tools for functional ecology. 2014.

gen3sis: engine for eco-evolutionary biodiversity modelling

1066 149. Cabral JS, Schurr FM. Estimating demographic models for the range dynamics of
1067 plant species. *Global Ecol Biogeogr.* 2010;19(1):85-97. doi: 10.1111/j.1466-
1068 8238.2009.00492.x.

1069 150. James S. Clark. Why trees migrate so fast: Confronting theory with dispersal
1070 biology and the paleorecord. *The American Naturalist.* 1998;152(2):204-24. doi:
1071 10.1086/286162.

1072 151. Quintero I, Wiens JJ. Rates of projected climate change dramatically exceed past
1073 rates of climatic niche evolution among vertebrate species. *Ecol Lett.* 2013;16(8):1095-
1074 103. Epub 2013/06/27. doi: 10.1111/ele.12144.

1075 152. Annan JD, Hargreaves JC. A new global reconstruction of temperature changes at
1076 the Last Glacial Maximum. *Climate of the Past.* 2012;9(1):367-76. doi: 10.5194/cp-9-
1077 367-2013.

1078 153. Scotese CR. Some thoughts on global climate change: the transition from
1079 icehouse to hothouse. *Paleomap project.* 2015;21:1 (2).

1080 154. Boucot AJ, Xu C, Scotese CR, Morley RJ. Phanerozoic paleoclimate: an atlas of
1081 lithologic indicators of climate. Tulsa, U.S.A.: Society of Economic Paleontologists and
1082 Mineralogists (Society for Sedimentary Geology); 2013.

1083 155. Spicer RA. Phytopaleoaltimetry: using plant fossils to measure past land surface
1084 elevation. In: Carina Hoorn, Allison Perrigo, Antonelli A, editors. *Mountains, Climate and*
1085 *Biodiversity* Oxford: Wiley. Oxford: Wiley; 2018. p. 95-109.

1086 156. Warren BH, Hagen O, Gerber F, Thebaud C, Paradis E, Conti E. Evaluating
1087 alternative explanations for an association of extinction risk and evolutionary
1088 uniqueness in multiple insular lineages. *Evolution.* 2018;72(10):2005-24. Epub
1089 2018/08/29. doi: 10.1111/evo.13582.

1090

1091

1092 Supporting Information captions

1093 *Animations*

1094 **Animation S1** *Reconstructed dynamic landscape L1 (i.e. world 65 Ma) with the*
1095 *environmental values used for the main case study.*

1096 **Animation S2** *Reconstructed dynamic landscape L2 (i.e. world 65 Ma) with the*
1097 *environmental values used for the main case study.*

1098 **Animation S3** *Theoretical dynamic landscape (i.e. theoretical island) with the environmental*
1099 *values used for the supplementary case study.*

1100 **Animation S4** *Dynamic simulated biodiversity patterns (i.e. M2 L1 world from 65 Ma to the*
1101 *present). The map shows the α diversity and the top and right graphs indicate the richness*
1102 *profile of longitude and latitude, respectively.*

1103 *Figures*

1104 **Figure S1** *Divergence increase per time-step d_i against the normalized occupied niche of*
1105 *isolated populations for models (A) M0 and M2, which assume temperature-independent*
1106 *divergence; and (B) M1, which assumes temperature-dependent divergence, where*
1107 *divergence relates to the mean of the realized temperature with three different d_{power}*
1108 *values.*

1109 **Figure S2** *Non-exhaustive probability density functions of the explored dispersal parameters*
1110 *in a Weibull distribution with shape ϕ of 1, 2 and 5 and Ψ of 550, 650, 750 and 850.*

1111 **Figure S3** *Frequencies of simulated normalized LDG slope (histogram) with empirical LDG*
1112 *for four main groups (dashed grey line) and acceptance range (red line). Frequencies for*

gen3sis: engine for eco-evolutionary biodiversity modelling

1113 *models (A) M0, (B) M1, (C) M2 with total frequency and frequency discriminated for each*
1114 *landscape, i.e. L1 and L2.*

1115 **Figure S4** *Correlation of model parameters and three emerging patterns for all models and*
1116 *landscapes (A) M0 L1, (B) M0 L2, (C) M1 L1, (D) M1 L2, (E) M2 L1, and (F) M2 L2.*

1117 *Emerging patterns: (i) phylogeny beta is the phylogenetic tree imbalance statistic measured*
1118 *as the value that maximizes the likelihood in the β -splitting model; (ii) range quant 0.95% is*
1119 *the value of the 95% quantile of the species range area distribution and; (iii) LDG slope is*
1120 *the slope of the linear regression of species richness.*

1121 **Figure S5** *Effects of grid cell size on simulations of M2 L1. (A) Correlation of grid cell, LDG*
1122 *slope and other summary statistics. (B) Simulated LDG slope and grid cell size, showing a*
1123 *significant effect of spatial resolution on LDG slope.*

1124 **Figure S6** *Normalized richness of (A) selected simulation, (B) terrestrial mammals, (C) birds,*
1125 *(D) amphibians and (E) reptiles, with Pearson correlation values for comparisons between*
1126 *simulated and empirical data.*

1127 **Figure S7** *Results of the island case study showing (A) landscape size and environmental*
1128 *dynamics and (B) results of three experiments (i.e. lower, equal and higher trait evolution*
1129 *compared with the temporal environmental variation). The time series in (B) shows γ*
1130 *richness (log10 scale) on theoretical oceanic islands, following the geomorphological*
1131 *dynamics of islands. Thick lines indicate the average of the replicates, whereas thin lines*
1132 *indicate SD envelopes (n=30 for each trait evolutionary rate scenario). The dashed grey*
1133 *vertical bar crossing the entire plot indicates the period in which the island reaches its*
1134 *maximum size.*

gen3sis: engine for eco-evolutionary biodiversity modelling

1135 *Notes*

1136 **Note S1** *Global case study: emergence of the LDG from environmental changes of the*

1137 *Cenozoic.*

1138 **Note S2** *Island case study: does trait evolution impact biodiversity dynamics?*

1139 **Note S3** *Gen3sis pseudo-code.*