Received: 1 March 2022 | Revised: 30 September 2022 | Accepted: 4 October 2022

DOI: 10.1111/geb.13604

Global Ecology and Biogeography A Journal of Macroecology

WILEY

REVIEW ARTICLE

Towards a genetic theory of island biogeography: Inferring processes from multidimensional community-scale data

Isaac Overcast[1,2] | Guillaume Achaz[3,4,5] | Robin Aguilée[6] | Carmelo Andújar[7] | Paula Arribas[7] | Thomas J. Creedy[8,9] | Evan P. Economo[10,11] | Rampal S. Etienne[12] | Rosemary Gillespie[13] | Claire Jacquet[14,15] | Flora Jay[16] | Susan Kennedy[10] | Henrik Krehenwinkel[17] | Amaury Lambert[1,18] | Emmanouil Meramveliotakis[19] | Víctor Noguerales[7,19] | Benoit Perez-Lamarque[1] | George Roderick[13] | Haldre Rogers[20] | Megan Ruffley[21] | Isabel Sanmartin[22] | Alfried P. Vogler[8,9] | Anna Papadopoulou[19] | Brent C. Emerson[7] | Hélène Morlon[1]

[1]Institut de Biologie de l'ENS (IBENS), Ecole Normale Supérieure, CNRS, INSERM, Université PSL, Paris, France

[2]Department of Vertebrate Zoology, American Museum of Natural History, New York, New York, USA

[3]Institut Systématique Evolution Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, Paris, France

[4]SMILE Group, CIRB, UMR 7241, Collège de France, CNRS, INSERM, Paris, France

[5]Éco-anthropologie, Muséum National d'Histoire Naturelle, CNRS UMR 7206, Université de Paris, Paris, France

[6]UMR5174 (Laboratoire Evolution et Diversité Biologique), CNRS, Université Toulouse III Paul Sabatier, Toulouse, France

[7]Island Ecology and Evolution Research Group, Instituto de Productos Naturales y Agrobiología, CSIC (IPNA-CSIC), Santa Cruz de Tenerife, Spain

[8]Department of Life Sciences, Natural History Museum, London, UK

[9]Department of Life Sciences, Imperial College London, Ascot, UK

[10]Biodiversity and Biocomplexity Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan

[11]Radcliffe Institute for Advanced Study, Harvard University, Cambridge, Massachusetts, USA

[12]Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands

[13]Department of Environmental Science, Policy, and Management, University of California, Berkeley, California, USA

[14]Department of Aquatic Ecology, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

[15]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zürich, Switzerland

[16]Université Paris-Saclay, CNRS UMR9015, Inria, Laboratoire Interdisciplinaire des Sciences du Numérique, Orsay, France

[17]Department of Biogeography, Trier University, Trier, Germany

[18]Center for Interdisciplinary Research in Biology (CIRB), Collège de France, CNRS UMR 7241, INSERM U 1050, Université PSL, Paris, France

[19]Department of Biological Sciences, University of Cyprus, Nicosia, Cyprus

[20]Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, Iowa, USA

[21]Carnegie Institution for Science, Stanford, California, USA

[22]Real Jardín Botánico (RJB), CSIC, Madrid, Spain

Correspondence
Isaac Overcast, Institut de Biologie de l'ENS (IBENS), Ecole Normale Supérieure, CNRS, INSERM, Université PSL, Paris 75005, France.
Email: overcast@biologie.ens.fr

Abstract

Background: MacArthur and Wilson's theory of island biogeography has been a foundation for obtaining testable predictions from models of community assembly and for developing models that integrate across scales and disciplines. Historically, however,

Funding information
Horizon 2020 Framework Programme, Grant/Award Number: 810729

Handling Editor: Jose Eduardo Meireles

these developments have focused on integration across ecological and macroevolutionary scales and on predicting patterns of species richness, abundance distributions, trait data and/or phylogenies. The distribution of genetic variation across species within a community is an emerging pattern that contains signatures of past population histories, which might provide an historical lens for the study of contemporary communities. As intraspecific genetic diversity data become increasingly available at the scale of entire communities, there is an opportunity to integrate microevolutionary processes into our models, moving towards development of a genetic theory of island biogeography.

Motivation/goal: We aim to promote the development of process-based biodiversity models that predict community genetic diversity patterns together with other community-scale patterns. To this end, we review models of ecological, microevolutionary and macroevolutionary processes that are best suited to the creation of unified models, and the patterns that these predict. We then discuss ongoing and potential future efforts to unify models operating at different organizational levels, with the goal of predicting multidimensional community-scale data including a genetic component.

Main conclusions: Our review of the literature shows that despite recent efforts, further methodological developments are needed, not only to incorporate the genetic component into existing island biogeography models, but also to unify processes across scales of biological organization. To catalyse these developments, we outline two potential ways forward, adopting either a top-down or a bottom-up approach. Finally, we highlight key ecological and evolutionary questions that might be addressed by unified models including a genetic component and establish hypotheses about how processes across scales might impact patterns of community genetic diversity.

KEYWORDS
biodiversity dynamics, community ecology, community genetic diversity, metabarcoding, metagenomics, population genetics, stochastic biodiversity models

1 | INTRODUCTION

General interest in understanding the processes that have given rise to current patterns of biodiversity has spurred the development of diverse models to explain these patterns. One of the more influential is MacArthur and Wilson's equilibrium theory of island biogeography (ETIB; MacArthur & Wilson, 1963, 1967). Although originally developed to understand biodiversity patterns on islands, the ETIB has been used more generally to predict community patterns, thanks to the analogy between island and community assemblages, both of which result from limited dispersal from a larger species pool. The ETIB has catalysed >50 years of research on models that represent how species from regional or global pools (i.e., the metacommunity) assemble into local communities (Chesson, 2000; Hubbell, 2001; Leibold & Chase, 2017; Tilman, 2004; Vellend, 2016). These community assembly models provide explicit predictions about patterns that can be measured in the field, such as species abundance distributions (SADs; McGill et al., 2007), spatial turnover (Nekola & White, 1999), temporal turnover (Magurran et al., 2010) and species–area relationships (Lomolino, 2000). The availability of other types of data has spurred efforts to develop models that also generate community-scale predictions for such data, including species traits (Jacquet et al., 2017; Xu et al., 2020), interaction networks (Gravel et al., 2011; Thébault & Fontaine, 2010) and phylogenetic trees (Jabot & Chave, 2009; Morlon et al., 2011; Sanmartín et al., 2008; Valente et al., 2015). These models often require integration across disciplines to represent processes operating over different organizational levels (individuals/genotypes/species) and spatial scales, within the community or within the metacommunity, such as environmental filtering, interspecific interactions, trait evolution, speciation and extinction. Such unified models are challenging to build, but they allow a powerful inference of multiple processes potentially operating at and interacting across various scales.

One dimension of biodiversity that remains largely absent from predictions of such integrative models is the distribution of population genetic variation across species within communities (Ellegren & Galtier, 2016). This shortcoming means that community-level data for genetic diversity cannot be used to reinforce inferences of ecological and evolutionary processes using island biodiversity models. This is unfortunate, because community-scale genetic data provide a record of the formation of ecological communities on a population-genetic (intermediate) time-scale (on the order of tens to tens of thousands of years), which can complement information obtained from other community-scale patterns (Overcast et al., 2020). For example, the strength of the correlation between species abundance and genetic diversity might provide an indication of community age, because all species in young communities (e.g., those on recently formed volcanic islands) should have low genetic diversity (but see Caujapé-Castells et al., 2017), whereas abundant species from old communities (e.g., those on relatively ancient volcanic islands) might have accumulated more genetic diversity than rare species (Overcast et al., 2020). Genetic diversity might also capture signatures of historical changes in population size and cycles of genetic connectivity that drive incipient species into and through the "grey zone" of speciation (Roux et al., 2016; Salces-Castellano et al., 2020). Importantly, in the context of considering processes across organizational levels and spatial scales, genetic diversity provides a snapshot of population history at a time-scale that can shed light on the link between microevolutionary, ecological and macroevolutionary processes, which encompass patterns of intraspecific allele frequencies, species abundances and phylogenetic relationships, respectively. Therefore, the development of unified models of biodiversity that provide predictions for community-scale genetic data, along with other community-level predictions, has the potential to provide insight into numerous key ecological and evolutionary questions (Box 1).

Data availability is no longer a major obstacle to the integration of microevolutionary processes in inferences from unified models of biodiversity. Obtaining genetic data at the community scale has long been hindered by logistical constraints, but metagenetic methods, including whole-organism community DNA metabarcoding (Creedy et al., 2021; Deiner et al., 2017), environmental DNA (Cordier et al., 2020; Taberlet et al., 2012) or mitochondrial metagenomics (Crampton-Platt et al., 2016; Gómez-Rodríguez et al., 2015, 2017), are increasingly removing this constraint. Conventionally, these approaches have been implemented to approximate community profiles at the species level, but their potential also to generate population-level genetic data for entire communities has recently been revealed (Andújar et al., 2021; Arribas et al., 2021; Elbrecht et al., 2018; Schloissnig et al., 2013; Tsuji et al., 2020; Turon et al., 2020). As advances in sequencing technology continue to produce higher yields, metagenetics is increasingly being used for the characterization and monitoring of biodiversity in marine, terrestrial and freshwater environments. As a consequence, community-level genetic data are accumulating in massive online repositories (e.g., GenBank; https://www.ncbi.nlm.nih.gov/genbank), such that it is becoming increasingly possible to study intraspecific genetic variation across multiple communities from regional to global scales (i.e., "macrogenetics"; Leigh et al., 2021; Miraldo et al., 2016). One focus, therefore, is to encourage development of models that will provide an explicit biogeographical context (Edwards et al., 2022) for macrogenetic studies by integrating community-scale intraspecific genetic variation into mechanistic models to understand and predict how biodiversity accumulates across spatial scales and levels of organization.

The field of population genetics has produced countless models to infer historical processes of individual (or closely related) species (e.g., Kingman, 1982; Wright, 1931). However, these models are not well adapted to analyse community-scale data, because their focal unit of study is individual species. Comparative phylogeographical models aggregate multiple individual-population genetic models for co-distributed species (Arbogast & Kenagy, 2001; Edwards et al., 2022), but they lack an ecological (community) context. There are a handful of precedent efforts to develop community-level models that include predictions for the genetic diversity of all constituent species. Johnson et al. (2000) derived a population genetic model of island biogeography that allowed estimation of rates of migration, extinction and speciation and that made explicit predictions about how genetic divergence relates to island area and distance from the mainland (Ricklefs & Bermingham, 2004). Inspired by the unified neutral theory of biodiversity (UNTB; Hubbell, 2001), which is an individual-based neutral model of community assembly following ETIB, Laroche et al. (2015) developed a model that predicts the species genetic diversity correlation (SGDC), which measures the correlation between the genetic diversity of a focal taxon and species richness in the local community (Vellend, 2005). Empirical studies across a broad array of taxa have found either positive (Lamy et al., 2013; Papadopoulou et al., 2011) or negative genetic diversity–species richness correlations (Marchesini et al., 2018; Xu et al., 2016), consistent with predictions by Laroche et al. (2015). Finally, Overcast et al. (2020) developed the massive eco-evolutionary synthesis simulations (MESS) model, which combines Hubbell's UNTB with the neutral model of population genetics (Kimura, 1983). MESS provides predictions for the correlation between genetic diversity and abundance and for the species genetic distribution (SGD; Overcast et al., 2019), an analogue to the SAD, but where species abundance is replaced by a measure of species genetic diversity.

Given these few existing developments, the power of unified models and trends towards increased access to community-level population genetic data, we aim to promote the development of process-based biodiversity models that predict community genetic diversity patterns together with other community-scale patterns. For simplicity, we focus on genetic diversity patterns obtained from single-locus data, consistent with data that are typically generated within metabarcoding studies, but note that modelling of multilocus or metagenomic data is also possible and would only increase the resolution of model predictions. We review current models for studying biodiversity dynamics from the perspective of community ecology, microevolution and macroevolution, with a focus on their applicability to island-like systems. We highlight the strengths and weaknesses of these existing models, in terms of integrating across organizational levels and predicting realistic multidimensional community data including a genetic component. Finally, we discuss ongoing and potential future efforts to unify models operating at different organizational levels, in order to predict multidimensional community-scale data better.

BOX 1 Why are unified models including community-scale genetic variation important for understanding the processes that generate and maintain biodiversity?

Many open questions remain concerning how the biodiversity we observe in nature has accumulated. For example, how do the dynamics of adaptation, competition, speciation and colonization result in the communities that we observe today? And how have local environmental conditions and human-mediated disturbance modulated these processes? These questions can be addressed, in part, without unified models; for example, community ecology studies often suppose that the dynamics of community assembly play out on time-scales so short that speciation can be ignored, and macroevolutionary investigations rarely consider processes such as mutation and the accumulation of genetic diversity. The main value of unified models is precisely to facilitate investigation of feedback between such processes (i.e., addressing questions such as: how does *in situ* speciation reshape communities locally, and how does it impact the genetic diversity of the two descendant species and of all other species in the community?). A growing awareness of the importance of such feedback has bolstered interest in unified models (Cavender-Bares & Wilczek, 2003; McGill, 2010; Pennell & Harmon, 2013). Unified models might also increase our ability to infer which, among a series of alternative eco-evolutionary scenarios, is most likely to have produced the data we observe, by simultaneously leveraging the complementary information contained in multiple axes of biodiversity data. For example, scenarios that are not distinguishable from species abundance distributions alone might be distinguishable when also including community genetic data. We identify here a non-exhaustive set of questions for which models predicting individual (or few) types of data have had some, albeit limited, success in resolving, but for which the development of unified models would provide further advances.

- *Factors modulating commonness and rarity within communities*. How much do colonization times and early establishment demographics determine which species are common and which are rare in contemporary communities (e.g., tropical forests in Hawaii; Craven et al., 2019)? Have abundance relationships been maintained over time, and have historical changes in abundance followed deterministic (niche partitioning) or stochastic (neutral assembly) processes (e.g., tropical forests in Panama; Wang et al., 2013)? Ecological models of community assembly predicting only contemporary species abundances are unable to answer such questions, but unified models including a population genetic component for predicting the past and present genetic diversity of contemporary species could do so.
- *Factors promoting species richness within communities*. How much does *in situ* speciation contribute to the diversity of local communities (e.g., Caribbean *Anolis* lizards; Losos & Schluter, 2000)? How much do the dynamics of speciation and extinction at the metacommunity and community scales versus ecological limits imposed by local environmental conditions (e.g., mediated by total abundance in the community) explain the diversity of local communities (e.g., Rabosky & Glor, 2010)? Standard community assembly models do not account for diversification dynamics, whereas standard macroevolutionary models consider ecological limits at the clade rather than community level. These processes could be untangled with unified models of lineage diversification with a spatial local community component to generate nested local/global phylogenetic, richness and abundance patterns.
- *Impact of community context on genetic diversity*. Genetic diversity of individual species is the outcome of processes that take place within a community context, yet this context is rarely considered in population genetic models. For example, dispersal limitation between, competition (or other types of interactions) within, and adaptation to ecological conditions of islands or island-like habitat patches will all contribute to shape individual (and collective) species genetic diversity (e.g., Malagasy herpetofauna; Kuhn et al., 2022). Unified models of colonization, abundance and competition/interaction that predict abundance and genetic diversity for all species can provide such a community context for population genetic models.
- *Diversification and genetic diversity*. How much do metacommunity diversification dynamics explain community-scale genetic diversity? How does *in situ* speciation impact the genetic diversity of descendant species, and how does this contribute to shaping genetic diversity across the entire community (e.g., Neotropical cichlids; Kautt et al., 2020)? Answering these questions requires a unified model of macroevolution, community assembly and population genetics.

2 | EXISTING MODELS OF COMMUNITY, GENETIC AND COMPARATIVE BIODIVERSITY DATA

A model that aims to incorporate predictions of genetic diversity, as an additional dimension of community-scale patterns, together with abundances, traits and/or network structure, species richness and phylogenetic patterns (Figure 1) needs to incorporate processes ranging from ecological (e.g., dispersal, environmental filtering, demographic processes modulated by environmental factors and intra- and interspecific interactions) to microevolutionary (e.g., mutation, genetic drift and selection) and macroevolutionary (e.g., speciation, extinction and phenotypic evolution) (Figure 2). We expect that these processes will modulate community-scale genetic diversity

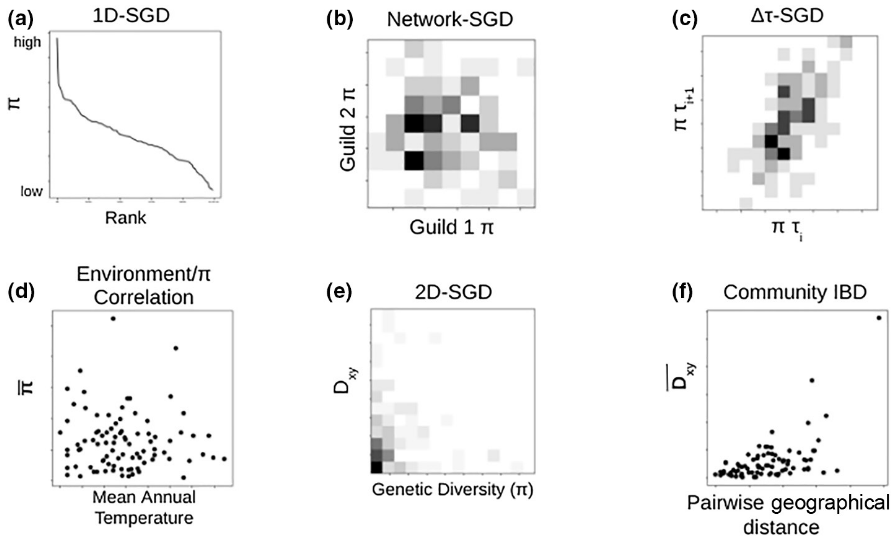

FIGURE 1 Patterns of community genetic diversity. Hypothetical illustrations of numerous patterns of single-locus community genetic diversity that can be envisioned within (a–d) and between (e, f) communities. Here, the focal summary statistic of the genetic data within communities is nucleotide diversity (π; Nei & Li, 1979), although other statistics, such as Watterson's estimator of θ (Watterson, 1975), Tajima's *D* (Tajima, 1989) or the site frequency spectrum (Wright, 1938), could provide complementary information. Between communities, we illustrate genetic divergence (D_{xy}; Nei & Li, 1979), but other population differentiation statistics, such as F_{ST} (Wright, 1965), could provide complementary information. (a) The one-dimensional species genetic diversity distribution (1D-SGD) can be represented as a rank ordering of π per species in the community. (b) The network-SGD is a pairwise histogram that summarizes π per species for two co-distributed interacting guilds. Darker bins in the histogram indicate higher numbers of interacting species pairs occupying this joint genetic diversity bin. (c) The $\Delta\tau$-SGD summarizes changes in genetic diversity through time given temporal samples from the same community. This is also a joint histogram, but with samples on the *x*-axis illustrating the π distribution at time τ_i and samples on the *y*-axis time τ_{i+1}. (d) The environment–$\overline{\pi}$ correlation depicts the relationship across local communities (each black point is a local community) between average π per species within each local community ($\overline{\pi}$) and a relevant environmental condition (here, mean annual temperature). (e) The two-dimensional species genetic diversity distribution (1D-SGD) summarizes the relationship between genetic diversity within species in the local community and genetic divergence between local and metacommunity sister-species pairs. It is a joint histogram depicting local π and genetic divergence (D_{xy}) from the metacommunity. (f) Isolation by distance (IBD) is a classic pattern in population genetics that represents the increase in genetic divergence between two populations as a function of the geographical distance separating them. A community-wide pattern of IBD can be formalized by calculating average D_{xy} ($\overline{D_{xy}}$) across all species shared between pairs of communities (black points) and plotting them as a function of the geographical distance separating the communities.

patterns (Box 2). If one is additionally interested in predicting spatially explicit patterns, similar to the species–area relationship (SAR; Durrett & Levin, 1996), the distance decay of similarity relationship (Morlon et al., 2008) or equivalent patterns for networks (Brose et al., 2004; Galiana et al., 2018), traits (Penone et al., 2016; Weinstein et al., 2014), genetic diversity (Baselga et al., 2015; Papadopoulou et al., 2011) and phylogenetic diversity (Morlon et al., 2011), then the model needs to incorporate spatially explicit dispersal and, potentially, environmental variation. Finally, prediction of temporal patterns requires following how patterns unfold through time under the chosen model (e.g., Missa et al., 2016). A rich suite of ecological, microevolutionary and macroevolutionary models that predict some of these patterns already exists. We begin with a short overview of some of these models, evaluating which patterns they are able to predict and whether the models provide realistic predictions, before moving into integration of genetic data into community-level models. We consider models that can be fitted to data using a variety of statistical techniques, ranging from likelihood-based to simulation-based inference (Box 3).

2.1 | Ecological models derived from the theory of island biogeography

The ETIB (MacArthur & Wilson, 1963, 1967) has been instrumental in ecology, inspiring many of the models that have subsequently been used in the field to predict community-scale data. Under the assumption that species are functionally equivalent (i.e., there are no interspecific differences in colonization or extinction rates), the theory predicts equilibrium richness as a product of immigration, which is dictated by distance from the mainland source, and extinction, which is affected by island size. Beyond predicting biodiversity patterns on islands, the ETIB has been used to predict community patterns, thanks to the analogy between continent–island and metacommunity–community systems (Hanski, 1998). For example, the ETIB naturally predicts SARs (Lomolino, 2000), and in general, these provide a reasonably good fit to empirical systems (Warren et al., 2015). Body-size distributions can be predicted by modelling body-size variation (e.g., the allometric theory of island biogeography; Jacquet et al., 2017),

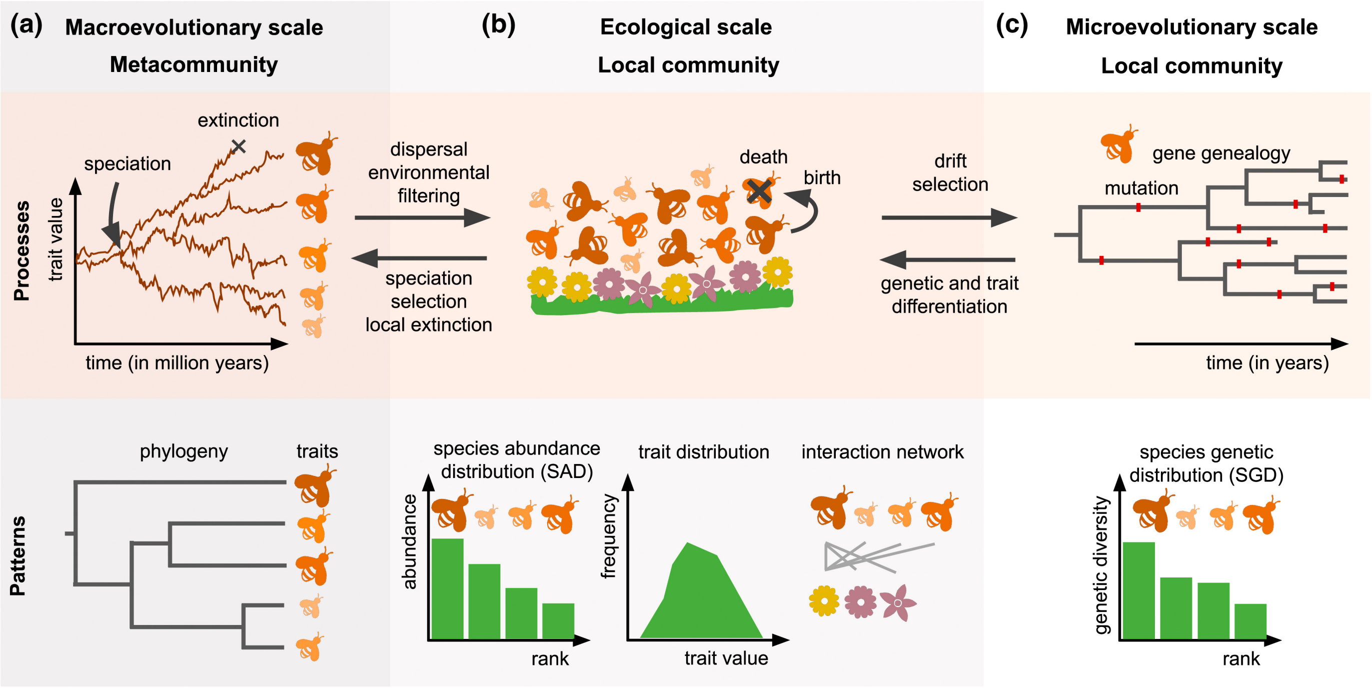

FIGURE 2 Processes across organizational levels required to model multidimensional biodiversity patterns that incorporate a genetic component. (a) At the macroevolutionary scale, speciation, extinction and trait evolution shape the phylogeny and traits of extant species in the metacommunity. Depicted species have different traits that are represented by colour and body size. (b) At the ecological scale, dispersal into and demographic processes within the local community, both of which are potentially modulated by trait values, shape local species abundance distributions (SADs) and trait distributions. In addition, interaction rules can shape the structure of interaction networks. Speciation, local extinction and trait selection can feed back on macroevolutionary processes. (c) At the microevolutionary scale, mutations accumulate, and selection, migration and fluctuations of species abundances shape species genetic diversity. At the scale of all species in the local community, patterns of community-scale genetic diversity [e.g., the species genetic distribution (SGD)] are predicted. Trait and genetic differentiation can feed back on both ecological and macroevolutionary processes.

under the assumption that island area and isolation will have differential impacts on colonization and extinction probability for large- versus small-bodied organisms. Food web properties, such as connectance (the fraction of realized interactions), can also be predicted (Gravel et al., 2011; Harmon et al., 2019). Several ETIB-inspired models can predict community-scale phylogenetic patterns, including DAISIE (Valente et al., 2015, 2020), which models immigration, extinction and speciation in an island/archipelago context, and DAMOCLES (Pigot & Etienne, 2015), which incorporates standard birth–death diversification processes at the regional scale with colonization and local extinction (but not speciation) processes at the local scale. The ETIB and its descendants are lineage based and, as such, they do not make predictions of abundance or genetic diversity. These models are also spatially implicit, in the sense that, although the rate of colonization can be a proxy for geographical distance, physical geography and landscape features are not modelled, hence they do not provide predictions of continuous spatial biodiversity patterns, such as the distance decay of similarity relationship.

The UNTB (Hubbell, 2001; Rosindell et al., 2011) is an individual-based model inspired by the ETIB, which models stochastic demographic processes in a small local community, within which all individuals of all species are functionally equivalent and which is at an equilibrium between local extinction and colonization from a large source pool of potential immigrants (metacommunity). The UNTB provides equilibrium predictions of species abundance distributions at local scales (Etienne, 2005, 2007) that are a good fit to temporally static SADs (Hubbell, 2001; Matthews & Whittaker, 2014) and can be relaxed to allow modelling of non-equilibrium abundance dynamics (Manceau et al., 2015; Missa et al., 2016; Overcast et al., 2019, 2020; Rosindell & Harmon, 2013), facilitating predictions of temporal patterns, such as changes in community composition through time (e.g., Kalyuzhny et al., 2015). Body-size distributions can be predicted within a UNTB-derived model by allowing variation in survival probability given body-size differences, while otherwise retaining ecological equivalence (e.g., O'Dwyer et al., 2009). Further extensions of the UNTB provide predictions for SADs on both sides of an interaction network, in addition to the structure of the network (Canard et al., 2012; Maliet et al., 2020). Spatially explicit temporal patterns of α- and β-diversity have also been modelled under static (Durrett & Levin, 1996; O'Dwyer & Green, 2010) or fluctuating (Gotelli et al., 2009; Jabot et al., 2020; Pontarp & Wiens, 2017; Rangel et al., 2007) environmental conditions, although how environmental stochasticity impacts spatial and temporal abundance patterns remains to be investigated fully. The UNTB also makes predictions of phylogenetic patterns at local

BOX 2 Hypothesized effects of biodiversity processes on community genetic diversity patterns

Community-scale patterns of genetic diversity reflect processes at all scales. Figure 3 illustrates examples of hypothesized effects of various processes on patterns described in Figure 1.

Figure 3a Diversification processes naturally impact genetic diversity. We can hypothesize that communities assembled from metacommunities within which the average speciation rate is high will have a reduced average genetic diversity at the community scale (continuous line). When descendant lineages are isolated by barriers to gene flow, and unless all segregating variation persists in both offspring lineages, there will be a net decrease of π in descendant lineages, relative to the ancestor. If the speciation rate outpaces the ability of offspring lineages to return to ancestral levels of genetic diversity, there will be a net decrease of average π at the community scale. We can also hypothesize that, in metacommunities subject to frequent extinction, the preferential extinction of genetically depauperate species (which have a reduced capacity for adaptation and increased potential for inbreeding depression) will elevate average π of the species remaining in the community (dashed line).

Figure 3b The nature of species interactions (e.g., antagonistic vs. mutualistic) impacts the structure of ecological networks, with potential cascading impacts on genetic diversity on both sides of the network. In mutualistic networks, few (abundant) generalists in general interact with many (rare) specialists (nested structure; Thébault & Fontaine, 2010), such that species abundance distributions [and therefore species genetic distributions (SGDs)] on each side of the network are expected to be more uneven, and abundances between interacting species more asymmetric (and therefore $\overline{\Delta\pi_{i,j}}$ larger) than in neutral or antagonistic networks.

Figure 3c The relative strength of selection versus ecological and genetic drift (neutral processes) modulates fluctuations in abundances that will influence fluctuations in genetic diversity across species in the community. Communities that change composition through time only as a result of drift will show large average $\Delta\tau$-SGD, owing to unconstrained fluctuations of abundances (and therefore π). Selection within the focal community will tend to constrain abundances, with well-adapted species obtaining and remaining at high abundance and with less well-adapted species existing primarily at low abundance or going locally extinct.

Figure 3d Differences in environmental conditions at different sampling sites will be correlated with community-scale genetic diversity to the extent that the environmental variable considered is biologically relevant. Under neutrality, species abundances are insensitive to the environment, fluctuating only as a function of drift, and therefore, average π in the community will not be correlated with environmental variables. In the presence of suitable environmental conditions, selection should reduce the local extinction rate and promote long-term persistence of species, increasing average π. In contrast, unsuitable conditions should act to increase demographic stochasticity, leading to elevated local extinction rates and reduced average π. The sensitivity of the community to any particular environmental variable will dictate the strength of its correlation with average π.

Figure 3e All things being equal, dispersal among communities should show a similar effect on community-scale genetic diversity to dispersal between structured populations of individual species, with increasing rates of migration increasing nucleotide diversity ($\overline{\pi}$; continuous line) within, and decreasing genetic distance (D_{xy}; dashed line) among, subdivided populations (Nei, 1973, 1987; Wright, 1965). However, variation in colonization time, effective population size and dispersal capacity among community members will cloud these signals. Investigating joint patterns of π and D_{xy} (e.g., the 2D-SGD) could shed light on both community history and variation in dispersal capacity. For example, controlling for effective population size, a relatively older, isolated species and a younger species connected by migration to the metacommunity might have similar π values, yet differ substantially regarding D_{xy}.

Figure 3f Community isolation by distance (IBD) should respond to dispersal in a similar way to classic (i.e., species-level) IBD, with low dispersal increasing average D_{xy} among sites with increasing geographical distance, increasing the slope of community IBD. With high dispersal, genetic divergence should erode more slowly with geographical distance, reducing the slope of community IBD. Variation in dispersal capacity among species within the community will cloud this signal, and therefore, accounting for this variation in calculating community IBD is an avenue for further study.

and metacommunity scales that fit the tree imbalance observed within empirical phylogenies reasonably well (Davies et al., 2011; Jabot & Chave, 2009). These have been improved by incorporation of protracted (Rosindell et al., 2010) or genealogical (Manceau et al., 2015; Rosindell et al., 2015) models of speciation. The UNTB has also served as a starting point for more general (including non-neutral) models of community structure (e.g., Haegeman & Etienne, 2017; Rosindell & Phillimore, 2011).

2.2 | Microevolutionary models for genetic diversity data

Population genetic models seek to represent the fundamental forces of evolution acting within species, including mutation, genetic drift, natural selection and migration, in the context of demographic history over time-scales of tens or hundreds to thousands of generations. They are typically applied to DNA sequence data from

BOX 3 Statistical approaches for fitting models of community assembly to multidimensional biodiversity data

Numerous statistical approaches exist for performing inference from real data, making trade-offs associated with model complexity, analytical and computational tractability, statistical power and accuracy, quantification of uncertainty, and interpretability.

When analytical likelihood formulas are available, these methods are a classical and powerful approach for fitting models to data. Likelihood-based approaches involve computing the probability of the data (e.g., a phylogeny or an abundance distribution) under a given model characterized by a set of parameters. In practice, likelihood methods do not always scale well to large datasets, in which case a composite likelihood approach approximating the full likelihood as a product of independent sub-likelihoods can be adopted (Varin et al., 2011). Alternatively, a data-augmentation approach (which consists of enhancing the data with unobserved events, such as extinctions) can simplify the computation of the likelihood and increase efficiency (Maliet & Morlon, 2020; Quintero & Landis, 2020). When used in combination with Bayesian statistics, such as Markov chain Monte Carlo (MCMC) sampling, likelihood-based methods naturally quantify uncertainty. Likelihood-based methods are the gold standard for most comparative phylogenetic methods (Morlon, 2014) and have also been used extensively to fit data to ecological (Rosindell et al., 2011) and microevolutionary (Kuhner et al., 1998; Pritchard et al., 2000) models.

For models of increasing complexity where likelihoods cannot be computed, a simulation-based approach can be adopted in an approximate Bayesian computation (ABC) framework (Beaumont, 2010; Csilléry et al., 2010). ABC approximates posterior model probabilities and parameter estimate distributions by comparing summary statistics calculated on the real data with identical summary statistics calculated on simulated data from tens of thousands or millions of simulations. The posteriors are constructed based on the proportion of simulated models and parameter values of a small fraction of retained simulations that are “closest” to the observed data. Classical ABC can be computationally demanding, because the number of simulations needed to sample parameter space grows exponentially with the number of parameters, although recent advances in combining machine learning with ABC have reduced this burden (Pudlo et al., 2015; Raynal et al., 2019). Sequential sampling methods (Beaumont et al., 2009; Lenormand et al., 2013) increase efficiency by iteratively sampling from regions of parameter space with the highest likelihood. ABC methods have seen broad adoption for numerous applications in population genetics (Beaumont, 2010). They has been used less frequently in ecological and macroevolutionary studies, particularly to infer model parameters (Jabot & Chave, 2009; Jabot & Lohier, 2016; Janzen et al., 2015; Overcast et al., 2019; Pontarp et al., 2019; Slater et al., 2012).

Machine learning (ML) methods offer a likelihood-free approach and are gaining significant attention for application to ecological and evolutionary questions (Schrider & Kern, 2018; Sheehan & Song, 2016). In the context of inference from multidimensional biodiversity data, we focus on supervised methods that allow both model selection and parameter estimation, unlike unsupervised “clustering” methods, such as principal component analysis. All supervised methods share a common process: (1) generate simulations under different models and parameter values; (2) train the ML on a subset of simulations; (3) test the ML model on held-out simulations to evaluate accuracy and recall; and (4) confront the ML with observed data to perform model classification and parameter estimation. Example applications of supervised ML include demographic model selection using population genetic summary statistics (e.g., Smith et al., 2017) and inference of historical population sizes using whole-genome data (e.g., Sanchez et al., 2020). ML has seen increasing use in macroevolutionary studies, for example to estimate phylogenies (Bhattacharjee & Bayzid, 2020) and for correlation of evolutionary rates (Tao et al., 2019). Early adoption of ML for ecological questions has been concentrated within the species distribution modelling literature (Gobeyn et al., 2019), with comparatively few applications in community ecology and with those being primarily focused on predicting microbial community structure (Thomas et al., 2018) and interaction patterns in ecological networks (Pichler et al., 2020). ML methods can be accurate and computationally efficient for inference under highly complex models. They avoid the curse of dimensionality (high-dimensional data are handled well), but suffer from reduced interpretability (but see Azodi et al., 2020) and difficulty in quantifying uncertainty around predictions, although this is an active area of research (Alaa & Van Der Schaar, 2020; Coulston et al., 2016).

homologous regions of DNA sampled from multiple individuals, but seldom in a community context. Individual-based population genetic models are typically constructed following either the forward-time classical Wright–Fisher (Wright, 1931) discrete time model of stochastic drift in finite populations or its backward-time coalescent (approximate) equivalent (Kingman, 1982; Tajima, 1983). Population genetic models tend to focus on individual neutrally evolving loci (e.g., Hudson, 2002; Kelleher et al., 2016), although effects of numerous types of selection can be included and are widely studied (see Bank et al., 2014) to predict non-neutral genetic diversity or the effect of drift and mutation on phenotypic trait variation within and among populations (Chakraborty & Nei, 1982; Lande, 1976). Population genetic models can also accommodate a spatial and ecological component to model demographic and adaptive histories of

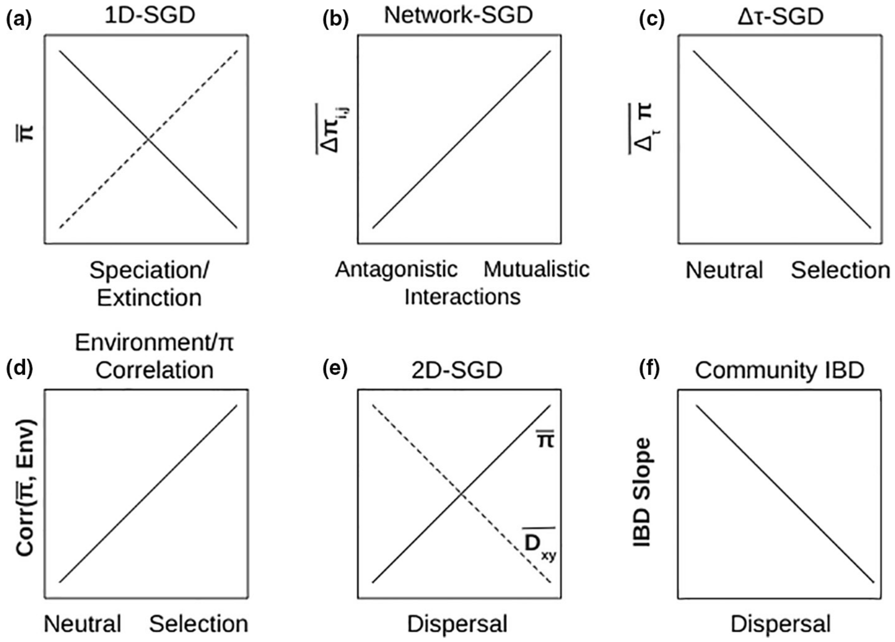

FIGURE 3 Examples of hypothesized effects of biodiversity processes on community genetic diversity patterns. The hypothesized effects are explained in Box 2. (a) The mean genetic diversity of species in the local community is denoted by $\overline{\pi}$. (b) The mean difference in π between all interacting species pairs, *i* and *j*, in the network is denoted by $\overline{\Delta\pi_{i,j}}$. (c) The mean difference in π values measured across species at two times separated by τ is denoted by $\overline{\Delta_\tau \pi}$. (d) The correlation coefficient of the environment–π correlation is denoted by Corr($\overline{\pi}$, Env). (e) The mean across species of genetic divergence between the local community and the metacommunity is denoted by $\overline{D_{xy}}$. (f) The slope of the relationship between $\overline{D_{xy}}$ across pairs of communities and the geographical distance separating them is denoted by the isolation-by-distance (IBD) slope. For other definitions, see legend to Figure 1.

arbitrary complexity. In the simplest scenarios, space is modelled implicitly, with isolation and migration models assuming connectivity between a small number of discrete panmictic populations (Wright, 1931), and genetic variation is modelled as a function of divergence times and migration rates among populations. Treating space explicitly opens the door for studying patterns such as isolation by distance (IBD; Wright, 1943), range expansion (Excoffier et al., 2008) and how intraspecific genetic diversity increases with area (e.g., Baselga et al., 2015). Furthermore, consideration of heterogeneous ecological landscapes that finely constrain migration and local adaptation gave rise to models integrating population genetics and landscape ecology (landscape genetics; Manel et al., 2003; Manel & Holderegger, 2013). One such contemporary spatially explicit coalescent approach entails a forward-in-time simulation of demography and migration, followed by a backward-in-time simulation of genealogies and genetic variation (Currat et al., 2019), allowing for modelling temporal patterns of genetic variation in a geographical context (e.g., Silva et al., 2018). Spatially explicit forward-in-time models based on Wright–Fisher processes are more computationally intensive than coalescent approaches, yet allow for much greater flexibility in modelling complex evolutionary scenarios incorporating fitness differences and selection on quantitative traits, allowing for modelling patterns of genetic variation, in addition to trait values across the landscape (e.g., Haller & Messer, 2019). In the past decade, models were also developed to predict population turnover and genetic offset based on ecological and genetic data (Fitzpatrick & Keller, 2015; Jay et al., 2012), akin to species distribution modelling, but at the intraspecific level. Nonetheless, population genetic models are predominantly designed to investigate processes within or among closely related species and, as such, they most often do not provide predictions for patterns above the species level (i.e., phylogenetic relationships or community-scale patterns of abundance).

Efforts to scale up microevolutionary models from the species level to multiple co-distributed species have occurred in comparative phylogeography (Arbogast & Kenagy, 2001; Edwards et al., 2022). Classic comparative phylogeographical studies attempted to infer histories of community assembly (e.g., Hewitt, 2000) from shared patterns of genetic variation across the landscape. The emergence of model-based comparative phylogeography (Hickerson et al., 2007; Huang et al., 2011; Xue & Hickerson, 2020) provided for explicit modelling of joint patterns of community-scale genetic variation under co-demographic processes, such as changes in population size (Gehara et al., 2017), while also accounting for vicariance and dispersal (Smith et al., 2014; Thom et al., 2020). These models implement a hierarchical, simulation-based approach, whereby coalescent simulations per focal taxon are parameterized conditional on hyperparameters that describe the assemblage-wide process (e.g., the proportion of co-expanding or co-diverging taxa). Although they describe patterns above the species level, comparative phylogeographical models lack an explicit ecological context and also do not consider or model phylogenetic relationships among co-distributed taxa.

2.3 | Macroevolutionary models for comparative data

The primary focus of macroevolutionary models is to represent the processes of diversification (speciation and extinction) and trait evolution for entire clades over time-scales of millions of years (Cornuault & Sanmartín, 2022; Morlon, 2014; Pennell & Harmon, 2013). They are typically applied to so-called "comparative data"; that is, the phylogenies and traits of extant species for taxonomically restricted groups of species. By focusing on taxonomically restricted groups of species, macroevolutionary models are not adapted a priori to modelling community-scale phylogenetic and trait patterns, yet they constitute an appropriate framework for communities arising from *in situ* radiations, as is the case for many radiations on isolated island archipelagos (Harmon et al., 2003; Mahler et al., 2010) or lakes (Janzen & Etienne, 2017).

Macroevolutionary models consider diversification and phenotypic evolution either separately (in this case, the processes are modelled hierarchically) or jointly (in this case, traits can influence diversification and vice versa; Maddison et al., 2007). The latter models, referred to as state-dependent speciation–extinction (SSE) models, have been reviewed elsewhere (Fitzjohn, 2012; Fitzjohn et al., 2009; Maddison & Fitzjohn, 2015). Recent advances in diversification modelling produce more realistic tree shapes by allowing variation in rate among lineages and through time (Maliet et al., 2019; Rabosky et al., 2013). Additional processes incorporated into lineage-based macroevolutionary models include protracted speciation (Etienne & Rosindell, 2012), clade-wise diversity-dependent speciation or extinction (Etienne et al., 2012; Mahler et al., 2010), inter- and/or intraspecific competition (Aristide & Morlon, 2019; Clarke et al., 2017; Drury et al., 2016; Xu et al., 2020), co-evolution across interacting lineages (Manceau et al., 2017) and the influence of environmental variation on macroevolutionary rates (Clavel & Morlon, 2017; Condamine et al., 2013). Some of these models include the modelling of dispersal events (Goldberg et al., 2011; Landis et al., 2013, 2021; Ree et al., 2005) and can be co-opted to make predictions about spatial patterns of biodiversity. They can potentially incorporate the effect of competition on trait evolution (Drury et al., 2018) or jointly on trait evolution and range occupancy (Quintero & Landis, 2020), allowing for the prediction of patterns of trait distribution and endemicity (for a more complete treatment of dispersal in macroevolutionary models, see Hackel & Sanmartín, 2021). These models incorporate more processes and can predict patterns that more closely match empirical ones, but they do not predict patterns other than the phylogenies and/or traits of extant taxa.

By construction, lineage-based models do not represent the individual-level processes of demography, mutation, genetic drift and both allele and trait selection that contribute to fluctuating species abundances and the accumulation of intraspecific genetic and trait diversity. The few individual-based models that also make predictions of macroevolutionary patterns include the UNTB, as already discussed above (Davies et al., 2011; Jabot & Chave, 2009; Missa et al., 2016), the model of Manceau et al. (2015), which is grounded in UNTB but relaxes both the zero-sum assumption and the point-mutation mode of speciation, and models that evolve species abundances and the traits of individuals along a phylogeny, which is either taken as data (Nuismer & Harmon, 2015) or simulated dynamically within the model (Duchen et al., 2020; McPeek, 2008). In principle, these latter models generate SADs while also modelling intraspecific genetic variation (Manceau et al., 2015) and inter- (McPeek, 2008) or intraspecific (Duchen et al., 2020; Nuismer & Harmon, 2015) trait variation; however, the predicted patterns have yet to be explored fully.

3 | TOWARDS A UNIFIED MULTIDIMENSIONAL COMMUNITY-SCALE THEORY OF ISLAND BIOGEOGRAPHY INCLUDING A GENETIC COMPONENT

To model multidimensional community data arising from processes across different organizational levels and spatial scales and including the genetic component, which was absent from many of the models described above, one must choose whether to model from the top down or bottom up. Top-down begins at the macroevolutionary scale and works down to the ecological and then microevolutionary scale (Figure 4). Bottom-up begins at the microevolutionary scale and naturally builds up to higher scales by letting patterns at all scales emerge from the underlying processes. Both strategies have advantages and disadvantages. Top-down models tend to scale well and can model clades of biologically meaningful phylogenetic diversity (hundreds of tips) but are, perhaps, less flexible in terms of modelling very complex processes. In addition, top-down models tend to be amenable to analytical likelihood expressions (Box 3), which makes them very computationally efficient. Bottom-up models tend to be based on individuals, which makes them costly to operate at biologically realistic scales, in particular up to the macroevolutionary scale, the trade-off being that they are more mechanistic (McGill, 2019). Any of these models can be fitted to multidimensional community-scale data to answer a variety of questions about the generation and maintenance of biological diversity (Box 1), using appropriate statistical inference techniques (Box 3).

3.1 | Top-down approaches

3.1.1 | From the macroevolutionary to the ecological scale

One top-down approach to combine macroevolutionary and ecological models consists of creating hybrid models that merge the metacommunity-scale dynamics of speciation and extinction with an ETIB-like community assembly process (Figure 4a). DAMOCLES (Pigot & Etienne, 2015), which incorporates standard birth–death diversification processes at the regional scale with colonization and local extinction processes at the local scale, is such a model.

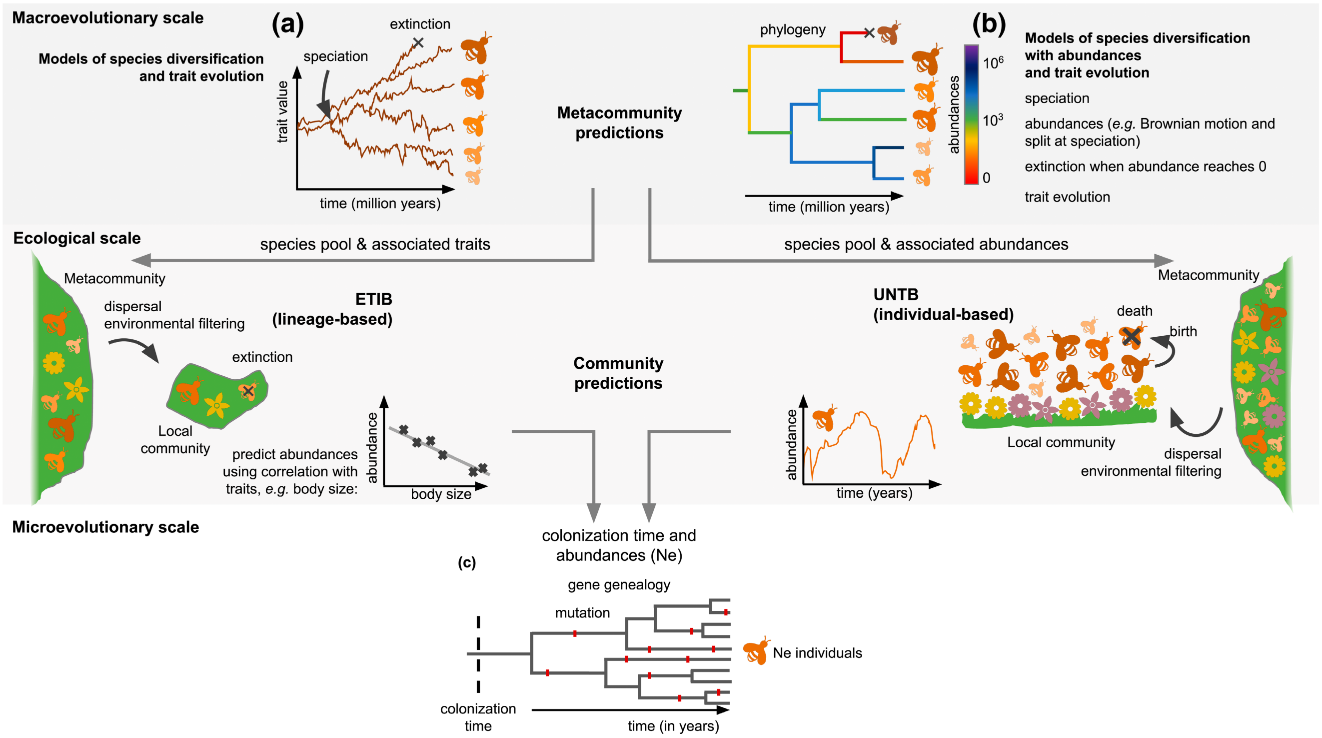

FIGURE 4 Top-down modelling approaches. Top-down modelling approaches linking macroevolutionary, ecological and microevolutionary scales can be formalized in two ways: (a) via a lineage-based approach using the equilibrium theory of island biogeography (ETIB) at the community scale; or (b) via an individual-based approach adopting the unified neutral theory of biodiversity (UNTB) at the community scale. (a) The lineage-based approach begins with a model of diversification and trait evolution, generating a global phylogeny and associated traits that constitute the source pool (metacommunity) for the ecological process. Dispersal to and extinction within the local community are mediated by the trait values of each lineage. Finally, abundances are predicted based on classical hypotheses regarding correlations between trait values (e.g., body size) and local abundance. (b) The individual-based approach begins by jointly modelling diversification and trait evolution, while additionally evolving species abundances along the branches. This constitutes the metacommunity source pool for ecological processes. Dispersal to the local community can be a function of abundance in the metacommunity and trait values. Fluctuations in abundances in the local community, including local extinction, emerge out of the individual-based birth–death–colonization process. Colonization times and historical changes in abundance per species are tracked throughout the process. (c) Both lineage-based and individual-based approaches produce predictions of colonization time and abundance, which can be used to parameterize microevolutionary models of coalescence and mutation for each species, resulting in patterns of community-scale genetic variation.

The spatial macroevolutionary models of diversification discussed above could also, in principle, serve this role, if they are configured to generate metacommunity–community patterns. For example, this could be achieved within the two-area GeoSSE model (Goldberg et al., 2011) by setting the dispersal rate from one of the areas (the community) to the other (the metacommunity) to zero. This would lead to a model representing processes very similar to those included in DAMOCLES, but with the additional advantage of representing speciation both within and between areas. Taking into consideration the metacommunity birth–death processes in DAISIE (Valente et al., 2015) would be an alternative approach to developing a similar hybrid model. These models illustrate the feasibility and usefulness of macroevolutionary–ecological models and pave the way for integrating more advanced ETIB-like community assembly processes (e.g., those accounting for interspecific interactions) within more advanced macroevolutionary models (e.g., those accounting for diversification rate heterogeneities, interspecific interactions and/or palaeoenvironmental fluctuations).

Unified neutral theory of biodiversity-like models can be used instead of ETIB-like models at the community scale as an alternative approach to creating hybrid models (Figure 4b), as implemented in MESS (Overcast et al., 2020). A key consideration for hybrid models is how to account for species abundances in macroevolutionary models in a principled way. In MESS, the metacommunity-scale phylogeny is generated with a birth–death lineage-based speciation process, and contemporary abundances are sampled from a logarithmic series distribution and assigned to species randomly. A more process-based approach would be to evolve abundances directly along phylogenetic branches as the lineage-based macroevolutionary process unfolds. In such a model, extinction events would naturally arise when species abundance falls below one, or any realistic minimum viable population size estimate. Speciation events would arise as the realization of a Poisson process, with rates following any of the models considered in macroevolution. Finally, there are several ways to evolve abundances. They could, for example, evolve as a Brownian process, with splits by random fission at each speciation event (Etienne & Haegeman, 2011; Hubbell & Lake, 2003). Alternatively, species-specific growth rates controlling changes in abundance could be evolved along branches as a Brownian process. The feedback from the ecological to the macroevolutionary scale might be realized by incorporating a protracted model of speciation (Rosindell et al., 2010), with the abundance structure of incipient lineages modulating the duration of speciation, whether complete speciation occurs and/or whether incipient species go extinct. Once a global phylogeny with abundances is generated by such a process, local community abundance distributions and phylogenetic relationships can be sampled following a classical UNTB individual-based simulation procedure.

From these neutral models, which assume that species are demographically and functionally equivalent, and following the same reasoning, several useful extensions are possible. In particular, existing macroevolutionary models of trait evolution could be used at the metacommunity scale to generate trait histories along phylogenetic branches, in addition to species-specific trait values at the tips, from which local community trait distributions can be sampled. In the non-neutral version of DAMOCLES, for example, traits influence colonization rates through either habitat filtering or competition with residents. In the non-neutral version of MESS, relaxation of the ecological neutrality assumption allows the influence of biotic and abiotic interaction to modulate individual survival. Although MESS has explored simple models of competition and environmental filtering, variation within an ecologically relevant trait (e.g., body size) could influence species abundances through more biologically realistic interspecific interactions, such as mutualism or facilitation. Recent lineage-based models of phenotypic trait evolution that model inter- and intraspecific variation jointly (Gaboriau et al., 2020; Kostikova et al., 2016) could also be included within macro-ETIB and macro-UNTB approaches, allowing for more fine-grained modelling of ecological interactions.

3.1.2 | From the ecological to the microevolutionary scale

After developing top-down models linking macroevolutionary and ecological processes, the next step is to extend these to the microevolutionary scale. To achieve this, one possibility is to constrain microevolutionary models per species to be contingent on divergence times and abundance histories predicted by the linked macroevolutionary–ecological model (e.g. Overcast et al., 2020). These constraints are justified by the observed correlation between census and effective size in natural populations (James & Eyre-Walker, 2020). This approach assumes that there is no post-divergence gene flow between sister species and that all lineages have sorted at the time of sampling. These simplifying assumptions are reasonable under low rates of diversification and/or small effective population sizes with respect to branching times. However, in cases where post-divergence gene flow cannot be ignored, species ages can be scaled to generation time, and migration rates among lineages can be modelled in a fully parameterized coalescent framework for the entire community.

This general approach can be applied to either macro-ETIB or macro-UNTB macroevolutionary–ecological models. As a prelude to this, Johnson et al. (2000) derived formulas based on the ETIB for average genetic divergence between island and mainland taxa as functions of time, island area and distance from the mainland. However, their model did not consider community-scale intraspecific diversity. ETIB models can naturally provide predictions for colonization times, but are not directly adapted to predicting fluctuations in species abundances required for genetic diversity predictions. The crux would be to use presupposed correlations between species abundances and characteristics of the species represented in these models. For example, the allometric theory of island biogeography (Jacquet et al., 2017) accounts for body-size differences across species and could be coupled with a documented relationship between body size and abundances (Damuth's law; Damuth, 1981) to provide community-level genetic diversity predictions at mutation–drift equilibrium. The trophic theory of island biogeography (Gravel et al., 2011) accounts for dietary breadth differences across species; by assuming that abundance scales with dietary breadth (i.e., generalists have greater abundance than specialists), community-level genetic diversity predictions at mutation–drift equilibrium could be obtained.

Macro-UNTB macroevolutionary–ecological models generate direct predictions of fluctuations in abundance within communities, which can be used to constrain microevolutionary models at the scale of individual species. Two such individual-based simulation approaches have been developed (Laroche et al., 2015; Overcast et al., 2019). Both are inspired by the UNTB, but they differ in the model used to predict genetic diversity. Laroche et al. (2015) paired a UNTB-inspired patch-based model with a forward-in-time mutational process for one focal taxon in the local community to predict the relationship between species diversity and intraspecific genetic diversity (the SGDC). This model could be generalized to predict genetic diversity for all the species in the local community, rather than for only an individual taxon. Overcast et al. (2019) developed gimmeSAD, which pairs a forward-time UNTB continent–island model tracking abundance and colonization time per species with a backward-in-time coalescent model per species to predict SADs and the community-level genetic diversity distribution jointly (SGD; Figure 2a). Another approach that suggests itself is to parameterize hierarchical comparative phylogeographical models directly using expectations from a macro-UNTB model. Indeed, lacking both a colonization and a community-assembly process, comparative phylogeographical models assume that all species have occupied the local landscape for the same amount of time and that the effective population sizes of individual species are sampled independently. Conditioning hyperparameter distributions on theoretical macro-UNTB expectations of species age and abundance would allow relaxation of these assumptions. Ultimately, these different approaches will result in a model representing the same processes.

These outlines of ETIB- and UNTB-based linked ecological and microevolutionary processes suggest numerous further avenues for development. For example, an eco-microevolutionary model of interactions among species of co-distributed guilds (e.g., plants–pollinators or hosts–parasites) could allow joint predictions of community genetic diversity and network structure, a possibility foreshadowed by recent work (Bunnefeld et al., 2018; Satler & Carstens, 2017; Stone et al., 2012). Linked spatially explicit models (Currat et al., 2019; Haller & Messer, 2019) would allow for ecologically informed predictions for novel community-scale spatial patterns of genetic diversity, such as the genetic diversity–area relationship, the correlation between genetic and species β-diversity, and the distance decay of genetic diversity (i.e., community-scale isolation by distance).

3.1.3 | From the macroevolutionary to the microevolutionary scale

There are existing conceptual frameworks that go directly from the macroevolutionary to the microevolutionary scale, without a community assembly component. Models that consider both interspecific and intraspecific divergences, such as the multispecies coalescent (MSC; Degnan & Rosenberg, 2009), the generalized mixed Yule coalescent (GMYC; Pons et al., 2006), the polymorphisms-aware phylogenetic model (POMO; De Maio et al., 2015) and automatic barcode gap discovery (ABGD; Puillandre et al., 2012), provide such a framework, although initially developed as species-delimitation or phylogenetic inference methods. Simulations under these frameworks model: (1) a species tree using typical Yule or birth–death macroevolutionary processes; (2) gene trees (MSC, POMO and ABGD) or intraspecific divergences (GMYC), using typical coalescent microevolutionary processes, with or without incomplete lineage sorting; and (3) potentially, lay down mutations on the gene trees. Obtaining community-scale patterns from these existing frameworks would require sampling species from the species tree, which could be done following either the macro-ETIB or macro-UNTB approach detailed above. Although this would lead to a model that converges to the macro–eco–micro model described above, it might provide additional possibilities. For example, pairing the multispecies coalescent with a mutational model underlying quantitative trait loci can generate phenotypic trait distributions in the presence of gene tree discordance (Mendes et al., 2018).

3.2 | Bottom-up unification approaches

A key obstacle for building bottom-up biodiversity models is to model speciation by the accumulation of genetic incompatibilities between populations (Orr, 1995). Among these microevolutionary models, those incorporating trait variation governed by an explicit genetic component can link demographic and speciation processes to differences in fitness within an ecological context and predict patterns above the species level (e.g., for a radiating clade; Gavrilets & Losos, 2009). Several such models have been proposed (e.g., Aguilée et al., 2018; Garwood et al., 2019; Gavrilets & Vose, 2005), all of which are individual-based birth–death processes that model genetic evolution via mutation and recombination within an ecological context. Including biotic interactions in such models is a pathway to more biologically relevant bottom-up models (e.g., Pontarp & Wiens, 2017; Thompson et al., 2020). Given that these models are primarily concerned with speciation, patterns of intraspecific genetic diversity are generated, but have not been evaluated. Additionally, these genetically explicit models tend to represent closed systems suitable for modelling *in situ* radiations, but lack a colonization process adapted to modelling the community structure of spatially isolated local communities (but see Aguilée et al., 2018; Gascuel et al., 2015, which model space explicitly). Opportunities remain for the development of models underpinned by a genetically explicit speciation process to make predictions about linked patterns of SADs, the distributions of traits and genetic variation, and phylogenetic and network structure.

Alternative attempts to model speciation from the accumulation of population genetic incompatibilities have been made using models either derived from or inspired by the UNTB. Instead of the traditional phenomenological models of speciation (e.g., point mutation and various derivations; Hubbell & Lake, 2003; Rosindell et al., 2010), speciation processes governed by an underlying process of genomic evolution and accumulation of incompatibility have been incorporated (e.g., de Aguiar et al., 2009; Hagen et al., 2021; Manceau et al., 2015; Marin et al., 2020; Melián et al., 2010). In these neutral models, speciation proceeds by the accumulation of genetic differentiation and is governed by a threshold effect (Gavrilets, 2004), whereby subpopulations that are differentiated by a given amount of genetic divergence are considered to be species. Although these models typically implement a fixed speciation threshold, sampling from a distribution of speciation thresholds per species would better align with observations of the "grey zone" of speciation in empirical systems (Roux et al., 2016). These models make joint predictions about community genetic diversity, abundance and phylogenetic structure (Costa et al., 2019) and, in some cases, spatial genetic patterns (Baptestini et al., 2013; de Aguiar et al., 2009), although these patterns have yet to be explored fully. These models are neutral with respect to genetics, hence there are neither fitness differences nor selection, and therefore, no direct effect of genetics on species abundance. Models incorporating both neutral and adaptive genetic components have been proposed (e.g., Aguilée et al., 2018; Gascuel et al., 2015), and investigating their predictions of community-scale genetic diversity and fitting these to empirical data will be fruitful avenues of future research.

4 | CONCLUSIONS

The distribution of genetic diversity, within and among species within ecological communities, is an emerging area of interest, but methodological development is needed to shed light on how ecological and evolutionary processes structure such diversity. Models of community-scale genetic diversity are in their infancy, with some limited but promising progress. Our focus in this perspective has been the modelling of single-locus data derived from metabarcoding studies, which are the predominant community-scale genetic datasets that are becoming widely obtainable. As sequencing throughput continues to increase, it should become possible to move beyond the generation of single-locus community genetic data to whole-community reduced representation sequencing datasets, consisting of thousands or tens of thousands of independent anonymous loci (Andrews et al., 2016), or even beyond, to community whole-genome data (e.g., Forster et al., 2016). For some model systems, such community-scale genomic data exist already (e.g., Neotropical cichlids; Kautt et al., 2020), providing opportunities to elucidate the role of genomic architecture in driving eco-evolutionary dynamics (Rudman et al., 2018). Population genetic models capturing complex genomic processes, including multilocus selection and fine-scale genome-wide mutation and recombination rate variation, allow for modelling these new types of genetic data (Haller & Messer, 2019; Kelleher et al., 2016). Integration of these models within the unifying frameworks highlighted here has the potential to provide even greater insights into the processes shaping biodiversity across scales. Our assessment of different potential approaches for building models that integrate processes across scales highlights that, despite the diversity of models developed in community ecology, population genetics and macroevolution, distinct envisioned efforts to expand their application across scales lead to convergent models. This leads us to conclude on the optimistic note that the long-sought, truly unified theory of biodiversity might be within close reach.

AUTHOR CONTRIBUTIONS

All authors contributed substantially to conceptual development. I.O. and H.M. wrote the first draft of the manuscript, and all authors contributed to manuscript revisions.

ACKNOWLEDGEMENTS

We thank the Morlon group for thoughtful and useful discussions. We acknowledge the support of the European Union's Horizon 2020 research and innovation programme through grant agreement number 810729.

FUNDING INFORMATION

This work is a product of the iBioGen project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 810729.

CONFLICT OF INTEREST

The authors declare no competing interests.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

ORCID

Isaac Overcast https://orcid.org/0000-0001-8614-6892
Robin Aguilée https://orcid.org/0000-0002-3388-9921
Claire Jacquet https://orcid.org/0000-0002-7913-0689
Benoit Perez-Lamarque https://orcid.org/0000-0001-7112-7197
Brent C. Emerson https://orcid.org/0000-0003-4067-9858
Hélène Morlon https://orcid.org/0000-0002-3195-7521

REFERENCES

Aguilée, R., Gascuel, F., Lambert, A., & Ferriere, R. (2018). Clade diversification dynamics and the biotic and abiotic controls of speciation and extinction rates. *Nature Communications*, *9*(1), 3013.

Alaa, A., & Van Der Schaar, M. (2020). Discriminative jackknife: Quantifying uncertainty in deep learning via higher-order influence functions. In H. D. Iii & A. Singh (Eds.), *Proceedings of the 37th International conference on machine learning* (Vol. *119*, pp. 165–174). PMLR.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews. Genetics*, *17*(2), 81–92.

Andújar, C., Creedy, T. J., Arribas, P., López, H., Salces-Castellano, A., Pérez-Delgado, A. J., Vogler, A. P., & Emerson, B. C. (2021). Validated removal of nuclear pseudogenes and sequencing artefacts from mitochondrial metabarcode data. *Molecular Ecology Resources*, *21*(6), 1772–1787.

Arbogast, B. S., & Kenagy, G. J. (2001). Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography*, *28*(7), 819–825.

Aristide, L., & Morlon, H. (2019). Understanding the effect of competition during evolutionary radiations: an integrated model of phenotypic and species diversification. *Ecology Letters*, *22*(12), 2006–2017.

Arribas, P., Andújar, C., & Salces-Castellano, A. (2021). The limited spatial scale of dispersal in soil arthropods revealed with whole-community haplotype-level metabarcoding. *Molecular Ecology*, *30*(1), 48–61.

Azodi, C. B., Tang, J., & Shiu, S.-H. (2020). Opening the black box: Interpretable machine learning for geneticists. *Trends in Genetics: TIG*, *36*(6), 442–455.

Bank, C., Ewing, G. B., Ferrer-Admettla, A., Foll, M., & Jensen, J. D. (2014). Thinking too positive? Revisiting current methods of population genetic selection inference. *Trends in Genetics: TIG*, *30*(12), 540–546.

Baptestini, E. M., de Aguiar, M. A. M., & Bar-Yam, Y. (2013). Conditions for neutral speciation via isolation by distance. *Journal of Theoretical Biology*, *335*, 51–56.

Baselga, A., Gómez-Rodríguez, C., & Vogler, A. P. (2015). Multi-hierarchical macroecology at species and genetic levels to discern neutral and non-neutral processes. *Global Ecology and Biogeography: A Journal of Macroecology*, *24*(8), 873–882.

Beaumont, M. A. (2010). Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics*, *41*, 397–406.

Beaumont, M. A., Cornuet, J.-M., Marin, J.-M., & Robert, C. P. (2009). Adaptive approximate Bayesian computation. *Biometrika*, *96*(4), 983–990.

Bhattacharjee, A., & Bayzid, M. S. (2020). Machine learning based imputation techniques for estimating phylogenetic trees from incomplete distance matrices. *BMC Genomics*, *21*(1), 497.

Brose, U., Ostling, A., Harrison, K., & Martinez, N. D. (2004). Unified spatial scaling of species and their trophic interactions. *Nature*, *428*(6979), 167–171.

Bunnefeld, L., Hearn, J., Stone, G. N., & Lohse, K. (2018). Whole-genome data reveal the complex history of a diverse ecological community.

Proceedings of the National Academy of Sciences of the United States of America, *115*(28), E6507–E6515.

Canard, E., Mouquet, N., Marescot, L., Gaston, K. J., Gravel, D., & Mouillot, D. (2012). Emergence of structural patterns in neutral trophic networks. *PLoS One*, *7*, e38295.

Caujapé-Castells, J., García-Verdugo, C., Marrero-Rodríguez, Á., Fernández-Palacios, J. M., Crawford, D. J., & Mort, M. E. (2017). Island ontogenies, syngameons, and the origins and evolution of genetic diversity in the Canarian endemic flora. *Perspectives in Plant Ecology, Evolution and Systematics*, *27*, 9–22.

Cavender-Bares, J., & Wilczek, A. (2003). Integrating micro- and macroevolutionary processes in community ecology. *Ecology*, *84*(3), 592–597.

Chakraborty, R., & Nei, M. (1982). Genetic differentiation of quantitative characters between populations or species: I. Mutation and random genetic drift. *Genetics Research*, *39*(3), 303–314.

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, *31*(1), 343–366.

Clarke, M., Thomas, G. H., & Freckleton, R. P. (2017). Trait evolution in adaptive radiations: Modeling and measuring interspecific competition on phylogenies. *The American Naturalist*, *189*(2), 121–137.

Clavel, J., & Morlon, H. (2017). Accelerated body size evolution during cold climatic periods in the Cenozoic. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(16), 4183–4188.

Condamine, F. L., Rolland, J., & Morlon, H. (2013). Macroevolutionary perspectives to environmental change. *Ecology Letters*, *16*, 72–85.

Cordier, T., Alonso-Sáez, L., Apothéloz-Perret-Gentil, L., Aylagas, E., Bohan, D. A., Bouchez, A., Chariton, A., Creer, S., Frühe, L., Keck, F., Keeley, N., Laroche, O., Leese, F., Pochon, X., Stoeck, T., Pawlowski, J., & Lanzén, A. (2020). Ecosystems monitoring powered by environmental genomics: A review of current strategies with an implementation roadmap. *Molecular Ecology*, *30*(13), 2937–2958.

Cornuault, J., & Sanmartín, I. (2022). A road map for phylogenetic models of species trees. *Molecular Phylogenetics and Evolution*, *173*, 107483.

Costa, C. L. N., Lemos-Costa, P., Marquitti, F. M. D., Fernandes, L. D., Ramos, M. F., Schneider, D. M., Martins, A. B., & de Aguiar, M. A. M. (2019). Signatures of microevolutionary processes in phylogenetic patterns. *Systematic Biology*, *68*(1), 131–144.

Coulston, J. W., Blinn, C. E., Thomas, V. A., & Wynne, R. H. (2016). Approximating prediction uncertainty for random forest regression models. *Photogrammetric Engineering & Remote Sensing*, *82*(3), 189–197.

Crampton-Platt, A., Yu, D. W., Zhou, X., & Vogler, A. P. (2016). Mitochondrial metagenomics: Letting the genes out of the bottle. *GigaScience*, *5*, 15.

Craven, D., Knight, T. M., Barton, K. E., Bialic-Murphy, L., & Chase, J. M. (2019). Dissecting macroecological and macroevolutionary patterns of forest biodiversity across the Hawaiian archipelago. *Proceedings of the National Academy of Sciences*, *116*(33), 16436–16441.

Creedy, T. J., Andújar, C., Meramveliotakis, E., Noguerales, V., Overcast, I., Papadopoulou, A., Morlon, H., Vogler, A. P., Emerson, B. C., & Arribas, P. (2021). Coming of age for COI metabarcoding of whole organism community DNA: Towards bioinformatic harmonisation. *Molecular Ecology Resources*, *22*(3), 847–861.

Csilléry, K., Blum, M. G. B., Gaggiotti, O. E., & François, O. (2010). Approximate Bayesian computation (ABC) in practice. *Trends in Ecology & Evolution*, *25*(7), 410–418.

Currat, M., Arenas, M., Quilodràn, C. S., Excoffier, L., & Ray, N. (2019). SPLATCHE3: Simulation of serial genetic data under spatially explicit evolutionary scenarios including long-distance dispersal. *Bioinformatics*, *35*(21), 4480–4483.

Damuth, J. (1981). Population density and body size in mammals. *Nature*, *290*(5808), 699–700.

Davies, T. J., Allen, A. P., Borda-de-Água, L., Regetz, J., & Melián, C. J. (2011). Neutral biodiversity theory can explain the imbalance of phylogenetic trees but not the tempo of their diversification. *Evolution; International Journal of Organic Evolution*, *65*(7), 1841–1850.

de Aguiar, M. A. M., Baranger, M., Baptestini, E. M., Kaufman, L., & Bar-Yam, Y. (2009). Global patterns of speciation and diversity. *Nature*, *460*(7253), 384–387.

De Maio, N., Schrempf, D., & Kosiol, C. (2015). PoMo: An allele frequency-based approach for species tree estimation. *Systematic Biology*, *64*(6), 1018–1031.

Degnan, J. H., & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, *24*(6), 332–340.

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., de Vere, N., Pfrender, M. E., & Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895.

Drury, J., Clavel, J., Manceau, M., & Morlon, H. (2016). Estimating the effect of competition on trait evolution using maximum likelihood inference. *Systematic Biology*, *65*(4), 700–710.

Drury, J. P., Grether, G. F., Garland, T., Jr., & Morlon, H. (2018). An assessment of phylogenetic tools for analyzing the interplay between interspecific interactions and phenotypic evolution. *Systematic Biology*, *67*(3), 413–427.

Duchen, P., Alfaro, M. L., Rolland, J., Salamin, N., & Silvestro, D. (2020). On the effect of asymmetrical trait inheritance on models of trait evolution. *Systematic Biology*, *70*(2), 376–388.

Durrett, R., & Levin, S. (1996). Spatial models for species-area curves. *Journal of Theoretical Biology*, *179*(2), 119–127.

Edwards, S. V., Robin, V. V., Ferrand, N., & Moritz, C. (2022). The evolution of comparative phylogeography: Putting the geography (and more) into comparative population genomics. *Genome Biology and Evolution*, *14*(1), evab176.

Elbrecht, V., Vamos, E. E., Steinke, D., & Leese, F. (2018). Estimating intraspecific genetic diversity from community DNA metabarcoding data. *PeerJ*, *6*, e4644.

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews. Genetics*, *17*(7), 422–433.

Etienne, R. S. (2005). A new sampling formula for neutral biodiversity. *Ecology Letters*, *8*(3), 253–260.

Etienne, R. S. (2007). A neutral sampling formula for multiple samples and an "exact" test of neutrality. *Ecology Letters*, *10*(7), 608–618.

Etienne, R. S., & Haegeman, B. (2011). The neutral theory of biodiversity with random fission speciation. *Theoretical Ecology*, *4*(1), 87–109.

Etienne, R. S., Haegeman, B., Stadler, T., Aze, T., Pearson, P. N., Purvis, A., & Phillimore, A. B. (2012). Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1732), 1300–1309.

Etienne, R. S., & Rosindell, J. (2012). Prolonging the past counteracts the pull of the present: protracted speciation can explain observed slowdowns in diversification. *Systematic Biology*, *61*(2), 204–213.

Excoffier, L., Foll, M., & Petit, R. J. (2008). Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 481–501.

Fitzjohn, R. G. (2012). Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution/British Ecological Society*, *3*(6), 1084–1092.

FitzJohn, R. G., Maddison, W. P., & Otto, S. P. (2009). Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Systematic Biology*, *58*(6), 595–611.

Fitzpatrick, M. C., & Keller, S. R. (2015). Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters*, *18*(1), 1–16.

Forster, S. C., Browne, H. P., Kumar, N., Hunt, M., Denise, H., Mitchell, A., Finn, R. D., & Lawley, T. D. (2016). HPMCD: The database of human microbial communities from metagenomic datasets and microbial reference genomes. *Nucleic Acids Research*, *44*(D1), D604–D609.

Gaboriau, T., Mendes, F. K., Joly, S., Silvestro, D., & Salamin, N. (2020). A multi-platform package for the analysis of intra- and interspecific trait evolution. *Methods in Ecology and Evolution*, *11*(11), 1439–1447.

Galiana, N., Lurgi, M., Claramunt-López, B., Fortin, M.-J., Leroux, S., Cazelles, K., Gravel, D., & Montoya, J. M. (2018). The spatial scaling of species interaction networks. *Nature Ecology & Evolution*, *2*(5), 782–790.

Garwood, R. J., Spencer, A. R. T., & Sutton, M. D. (2019). RE voSim: Organism-level simulation of macro and microevolution. *Palaeontology*, *62*(3), 339–355.

Gascuel, F., Ferrière, R., Aguilée, R., & Lambert, A. (2015). How ecology and landscape dynamics shape phylogenetic trees. *Systematic Biology*, *64*(4), 590–607.

Gavrilets, S. (2004). *Fitness landscapes and the origin of species (MPB-41)*. Princeton University Press.

Gavrilets, S., & Losos, J. B. (2009). Adaptive radiation: Contrasting theory with data. *Science*, *323*(5915), 732–737.

Gavrilets, S., & Vose, A. (2005). Dynamic patterns of adaptive radiation. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(50), 18040–18045.

Gehara, M., Garda, A. A., Werneck, F. P., Oliveira, E. F., da Fonseca, E. M., Camurugi, F., Magalhães, F. D. M., Lanna, F. M., Sites, J. W., Jr., Marques, R., Silveira-Filho, R., São Pedro, V. A., Colli, G. R., Costa, G. C., & Burbrink, F. T. (2017). Estimating synchronous demographic changes across populations using hABC and its application for a herpetological community from northeastern Brazil. *Molecular Ecology*, *26*(18), 4756–4771.

Gobeyn, S., Mouton, A. M., Cord, A. F., Kaim, A., Volk, M., & Goethals, P. L. M. (2019). Evolutionary algorithms for species distribution modelling: A review in the context of machine learning. *Ecological Modelling*, *392*, 179–195.

Goldberg, E. E., Lancaster, L. T., & Ree, R. H. (2011). Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology*, *60*(4), 451–465.

Gómez-Rodríguez, C., Crampton-Platt, A., Timmermans, M. J., Baselga, A., & Vogler, A. P. (2015). Validating the power of mitochondrial metagenomics for community ecology and phylogenetics of complex assemblages. *Methods in Ecology and Evolution*, *6*(8), 883–894.

Gómez-Rodríguez, C., Timmermans, M. J. T. N., Crampton-Platt, A., & Vogler, A. P. (2017). Intraspecific genetic variation in complex assemblages from mitochondrial metagenomics: Comparison with DNA barcodes. *Methods in Ecology and Evolution*, *8*(2), 248–256.

Gotelli, N. J., Anderson, M. J., Arita, H. T., Chao, A., Colwell, R. K., Connolly, S. R., Currie, D. J., Dunn, R. R., Graves, G. R., Green, J. L., Grytnes, J.-A., Jiang, Y.-H., Jetz, W., Kathleen Lyons, S., McCain, C. M., Magurran, A. E., Rahbek, C., Rangel, T. F. L. V. B., Soberón, J., … Willig, M. R. (2009). Patterns and causes of species richness: A general simulation model for macroecology. *Ecology Letters*, *12*(9), 873–886.

Gravel, D., Massol, F., Canard, E., Mouillot, D., & Mouquet, N. (2011). Trophic theory of island biogeography. *Ecology Letters*, *14*(10), 1010–1016.

Hackel, J., & Sanmartín, I. (2021). Modelling the tempo and mode of lineage dispersal. *Trends in Ecology & Evolution*, *36*(12), 1102–1112.

Haegeman, B., & Etienne, R. S. (2017). A general sampling formula for community structure data. *Methods in Ecology and Evolution*, *8*(11), 1506–1519.

Hagen, O., Flück, B., Fopp, F., Cabral, J. S., Hartig, F., Pontarp, M., Rangel, T. F., & Pellissier, L. (2021). gen3sis: A general engine for eco-evolutionary simulations of the processes that shape Earth's biodiversity. *PLoS Biology*, *19*(7), e3001340.

Haller, B. C., & Messer, P. W. (2019). SLiM 3: Forward genetic simulations beyond the Wright-Fisher model. *Molecular Biology and Evolution*, *36*(3), 632–637.

Hanski, I. (1998). Metapopulation dynamics. *Nature*, *396*(6706), 41–49.

Harmon, L. J., Andreazzi, C. S., Débarre, F., Drury, J., Goldberg, E. E., Martins, A. B., Melián, C. J., Narwani, A., Nuismer, S. L., Pennell, M. W., Rudman, S. M., Seehausen, O., Silvestro, D., Weber, M., & Matthews, B. (2019). Detecting the macroevolutionary signal of species interactions. *Journal of Evolutionary Biology*, *32*(8), 769–782.

Harmon, L. J., Schulte, J. A., 2nd, Larson, A., & Losos, J. B. (2003). Tempo and mode of evolutionary radiation in iguanian lizards. *Science*, *301*(5635), 961–964.

Hewitt, G. (2000). The genetic legacy of the quaternary ice ages. *Nature*, *405*(6789), 907–913.

Hickerson, M. J., Stahl, E., & Takebayashi, N. (2007). msBayes: Pipeline for testing comparative phylogeographic histories using hierarchical approximate Bayesian computation. *BMC Bioinformatics*, *8*, 268.

Huang, W., Takebayashi, N., Qi, Y., & Hickerson, M. J. (2011). MTML-msBayes: Approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. *BMC Bioinformatics*, *12*, 1.

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. Princeton University Press.

Hubbell, S. P., & Lake, J. K. (2003). The neutral theory of biodiversity and biogeography, and beyond. In T. Blackburn, & K. Gaston (Eds.), *Macroecology: Concepts and consequences* (pp. 45–63). Blackwell.

Hudson, R. R. (2002). Generating samples under a Wright–Fisher neutral model of genetic variation. *Bioinformatics*, *18*(2), 337–338.

Jabot, F., & Chave, J. (2009). Inferring the parameters of the neutral theory of biodiversity using phylogenetic information and implications for tropical forests. *Ecology Letters*, *12*(3), 239–248.

Jabot, F., Laroche, F., Massol, F., Arthaud, F., Crabot, J., Dubart, M., Blanchet, S., Munoz, F., David, P., & Datry, T. (2020). Assessing metacommunity processes through signatures in spatiotemporal turnover of community composition. *Ecology Letters*, *23*(9), 1330–1339.

Jabot, F., & Lohier, T. (2016). Non-random correlation of species dynamics in tropical tree communities. *Oikos*, *125*(12), 1733–1742.

Jacquet, C., Mouillot, D., Kulbicki, M., & Gravel, D. (2017). Extensions of island biogeography theory predict the scaling of functional trait composition with habitat area and isolation. *Ecology Letters*, *20*(2), 135–146.

James, J., & Eyre-Walker, A. (2020). Mitochondrial DNA sequence diversity in mammals: A correlation between the effective and census population sizes. *Genome Biology and Evolution*, *12*(12), 2441–2449.

Janzen, T., & Etienne, R. S. (2017). Inferring the role of habitat dynamics in driving diversification: Evidence for a species pump in Lake Tanganyika cichlids. *bioRxiv*, *2017*, 085431.

Janzen, T., Höhna, S., & Etienne, R. S. (2015). Approximate Bayesian computation of diversification rates from molecular phylogenies: Introducing a new efficient summary statistic, the nLTT. *Methods in Ecology and Evolution*, *6*(5), 566–575.

Jay, F., Manel, S., Alvarez, N., Durand, E. Y., Thuiller, W., Holderegger, R., Taberlet, P., & François, O. (2012). Forecasting changes in population genetic structure of alpine plants in response to global warming. *Molecular Ecology*, *21*(10), 2354–2368.

Johnson, K. P., Adler, F. R., & Cherry, J. L. (2000). Genetic and phylogenetic consequences of island biogeography. *Evolution; International Journal of Organic Evolution*, *54*(2), 387–396.

Kalyuzhny, M., Kadmon, R., & Shnerb, N. M. (2015). A neutral theory with environmental stochasticity explains static and dynamic properties of ecological communities. *Ecology Letters*, *18*(6), 572–580.

Kautt, A. F., Kratochwil, C. F., Nater, A., Machado-Schiaffino, G., Olave, M., Henning, F., Torres-Dowdall, J., Härer, A., Hulsey, C. D., Franchini, P., Pippel, M., Myers, E. W., & Meyer, A. (2020). Contrasting signatures of genomic divergence during sympatric speciation. *Nature*, *588*(7836), 106–111.

Kelleher, J., Etheridge, A. M., & McVean, G. (2016). Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLoS Computational Biology*, *12*(5), e1004842.

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge University Press.

Kingman, J. F. C. (1982). On the genealogy of large populations. *Journal of Applied Probability*, *19*(A), 27–43.

Kostikova, A., Silvestro, D., Pearman, P. B., & Salamin, N. (2016). Bridging inter- and intraspecific trait evolution with a hierarchical Bayesian approach. *Systematic Biology*, *65*(3), 417–431.

Kuhn, A., Gehara, M., Andrianarimalala, S. M., Rabibisoa, N., Randriamahatantsoa, B., Overcast, I., Raxworthy, C., Ruane, S., & Burbrink, F. T. (2022). Drivers of unique and asynchronous population dynamics in Malagasy herpetofauna. *Journal of Biogeography*, *49*(4), 600–616.

Kuhner, M. K., Yamato, J., & Felsenstein, J. (1998). Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics*, *149*(1), 429–434.

Lamy, T., Jarne, P., Laroche, F., Pointier, J.-P., Huth, G., Segard, A., & David, P. (2013). Variation in habitat connectivity generates positive correlations between species and genetic diversity in a metacommunity. *Molecular Ecology*, *22*(17), 4445–4456.

Lande, R. (1976). Natural selection and random genetic drift in phenotypic evolution. *Evolution; International Journal of Organic Evolution*, *30*(2), 314–334.

Landis, M. J., Eaton, D. A. R., Clement, W. L., Park, B., Spriggs, E. L., Sweeney, P. W., Edwards, E. J., & Donoghue, M. J. (2021). Joint phylogenetic estimation of geographic movements and biome shifts during the global diversification of viburnum. *Systematic Biology*, *70*(1), 67–85.

Landis, M. J., Matzke, N. J., Moore, B. R., & Huelsenbeck, J. P. (2013). Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology*, *62*(6), 789–804.

Laroche, F., Jarne, P., Lamy, T., David, P., & Massol, F. (2015). A neutral theory for interpreting correlations between species and genetic diversity in communities. *The American Naturalist*, *185*(1), 59–69.

Leibold, M. A., & Chase, J. M. (2017). *Metacommunity ecology*. Princeton University Press.

Leigh, D. M., van Rees, C. B., Millette, K. L., Breed, M. F., Schmidt, C., Bertola, L. D., Hand, B. K., Hunter, M. E., Jensen, E. L., Kershaw, F., Liggins, L., Luikart, G., Manel, S., Mergeay, J., Miller, J. M., Segelbacher, G., Hoban, S., & Paz-Vinas, I. (2021). Opportunities and challenges of macrogenetic studies. *Nature Reviews. Genetics*, *22*(12), 791–807.

Lenormand, M., Jabot, F., & Deffuant, G. (2013). Adaptive approximate Bayesian computation for complex models. *Computational Statistics*, *28*(6), 2777–2796.

Lomolino, M. V. (2000). Ecology's most general, yet protean pattern: The species-area relationship. *Journal of Biogeography*, *27*(1), 17–26.

Losos, J. B., & Schluter, D. (2000). Analysis of an evolutionary species–area relationship. *Nature*, *408*(6814), 847–850.

MacArthur, R. H., & Wilson, E. O. (1963). An equilibrium theory of insular zoogeography. *Evolution; International Journal of Organic Evolution*, *17*(4), 373–387.

MacArthur, R. H., & Wilson, E. O. (1967). *Theory of island biogeography. (MPB-1)*. Princeton University Press.

Maddison, W. P., & FitzJohn, R. G. (2015). The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic Biology*, *64*(1), 127–136.

Maddison, W. P., Midford, P. E., & Otto, S. P. (2007). Estimating a binary character's effect on speciation and extinction. *Systematic Biology*, *56*(5), 701–710.

Magurran, A. E., Baillie, S. R., Buckland, S. T., Dick, J. M., Elston, D. A., Scott, E. M., Smith, R. I., Somerfield, P. J., & Watt, A. D. (2010). Long-term datasets in biodiversity research and monitoring: Assessing change in ecological communities through time. *Trends in Ecology & Evolution*, *25*(10), 574–582.

Mahler, D. L., Revell, L. J., Glor, R. E., & Losos, J. B. (2010). Ecological opportunity and the rate of morphological evolution in the diversification of Greater Antillean anoles. *Evolution; International Journal of Organic Evolution*, *64*(9), 2731–2745.

Maliet, O., Hartig, F., & Morlon, H. (2019). A model with many small shifts for estimating species-specific diversification rates. *Nature Ecology & Evolution*, *3*(7), 1086–1092.

Maliet, O., Loeuille, N., & Morlon, H. (2020). An individual-based model for the eco-evolutionary emergence of bipartite interaction networks. *Ecology Letters*, *23*(11), 1623–1634.

Maliet, O., & Morlon, H. (2020). Fast and accurate estimation of species-specific diversification rates using data augmentation. *Systematic Biology*, *71*(2), 353–366.

Manceau, M., Lambert, A., & Morlon, H. (2015). Phylogenies support out-of-equilibrium models of biodiversity. *Ecology Letters*, *18*(4), 347–356.

Manceau, M., Lambert, A., & Morlon, H. (2017). A unifying comparative phylogenetic framework including traits coevolving across interacting lineages. *Systematic Biology*, *66*(4), 551–568.

Manel, S., & Holderegger, R. (2013). Ten years of landscape genetics. *Trends in Ecology & Evolution*, *28*(10), 614–621.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, *18*(4), 189–197.

Marchesini, A., Vernesi, C., Battisti, A., & Ficetola, G. F. (2018). Deciphering the drivers of negative species-genetic diversity correlation in Alpine amphibians. *Molecular Ecology*, *27*(23), 4916–4930.

Marin, J., Achaz, G., Crombach, A., & Lambert, A. (2020). The genomic view of diversification. *Journal of Evolutionary Biology*, *33*(10), 1387–1404.

Matthews, T. J., & Whittaker, R. J. (2014). Neutral theory and the species abundance distribution: Recent developments and prospects for unifying niche and neutral perspectives. *Ecology and Evolution*, *4*(11), 2263–2277.

McGill, B. J. (2010). Towards a unification of unified theories of biodiversity. *Ecology Letters*, *13*(5), 627–642.

McGill, B. J. (2019). The what, how and why of doing macroecology. *Global Ecology and Biogeography: A Journal of Macroecology*, *28*(1), 6–17.

McGill, B. J., Etienne, R. S., Gray, J. S., Alonso, D., Anderson, M. J., Benecha, H. K., Dornelas, M., Enquist, B. J., Green, J. L., He, F., Hurlbert, A. H., Magurran, A. E., Marquet, P. A., Maurer, B. A., Ostling, A., Soykan, C. U., Ugland, K. I., & White, E. P. (2007). Species abundance distributions: Moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters*, *10*(10), 995–1015.

McPeek, M. A. (2008). The ecological dynamics of clade diversification and community assembly. *The American Naturalist*, *172*(6), E270–E284.

Melián, C. J., Alonso, D., Vázquez, D. P., Regetz, J., & Allesina, S. (2010). Frequency-dependent selection predicts patterns of radiations and biodiversity. *PLoS Computational Biology*, *6*(8), e1000892.

Mendes, F. K., Fuentes-González, J. A., Schraiber, J. G., & Hahn, M. W. (2018). A multispecies coalescent model for quantitative traits. *eLife*, *7*, e36482.

Miraldo, A., Li, S., Borregaard, M. K., Flórez-Rodríguez, A., Gopalakrishnan, S., Rizvanovic, M., Wang, Z., Rahbek, C., Marske, K. A., & Nogués-Bravo, D. (2016). An Anthropocene map of genetic diversity. *Science*, *353*(6307), 1532–1535.

Missa, O., Dytham, C., & Morlon, H. (2016). Understanding how biodiversity unfolds through time under neutral theory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1691), 20150226.

Morlon, H. (2014). Phylogenetic approaches for studying diversification. *Ecology Letters*, *17*(4), 508–525.

Morlon, H., Chuyong, G., Condit, R., Hubbell, S., Kenfack, D., Thomas, D., Valencia, R., & Green, J. L. (2008). A general framework for the distance-decay of similarity in ecological communities. *Ecology Letters*, *11*(9), 904–917.

Morlon, H., Schwilk, D. W., Bryant, J. A., Marquet, P. A., Rebelo, A. G., Tauss, C., Bohannan, B. J. M., & Green, J. L. (2011). Spatial patterns of phylogenetic diversity. *Ecology Letters*, *14*(2), 141–149.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, *70*(12), 3321–3323.

Nei, M. (1987). *Molecular evolutionary genetics*. Columbia University Press.

Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(10), 5269–5273.

Nekola, J. C., & White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, *26*(4), 867–878.

Nuismer, S. L., & Harmon, L. J. (2015). Predicting rates of interspecific interaction from phylogenetic trees. *Ecology Letters*, *18*(1), 17–27.

O'Dwyer, J. P., & Green, J. L. (2010). Field theory for biogeography: A spatially explicit model for predicting patterns of biodiversity. *Ecology Letters*, *13*(1), 87–95.

O'Dwyer, J. P., Lake, J. K., Ostling, A., Savage, V. M., & Green, J. L. (2009). An integrative framework for stochastic, size-structured community assembly. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(15), 6170–6175.

Orr, H. A. (1995). The population genetics of speciation: The evolution of hybrid incompatibilities. *Genetics*, *139*(4), 1805–1813.

Overcast, I., Emerson, B. C., & Hickerson, M. J. (2019). An integrated model of population genetics and community ecology. *Journal of Biogeography*, *46*(4), 816–829.

Overcast, I., Ruffley, M., Rosindell, J., Harmon, L., Borges, P. A. V., Emerson, B. C., Etienne, R. S., Gillespie, R., Krehenwinkel, H., Luke Mahler, D., Massol, F., Parent, C. E., Patiño, J., Peter, B., Week, B., Wagner, C., Hickerson, M. J., & Rominger, A. (2020). A unified model of species abundance, genetic diversity, and functional diversity reveals the mechanisms structuring ecological communities. *Molecular Ecology Resources*, *21*(8), 2782–2800.

Papadopoulou, A., Anastasiou, I., Spagopoulou, F., Stalimerou, M., Terzopoulou, S., Legakis, A., & Vogler, A. P. Associate Editor: Diane S. Srivastava, & Editor: Judith L. Bronstein. (2011). Testing the species–genetic diversity correlation in the aegean archipelago: Toward a haplotype-based macroecology? *The American Naturalist*, *178*(2), 241–255.

Pennell, M. W., & Harmon, L. J. (2013). An integrative view of phylogenetic comparative methods: Connections to population genetics, community ecology, and paleobiology. *Annals of the New York Academy of Sciences*, *1289*, 90–105.

Penone, C., Weinstein, B. G., Graham, C. H., Brooks, T. M., Rondinini, C., Hedges, S. B., Davidson, A. D., & Costa, G. C. (2016). Global mammal beta diversity shows parallel assemblage structure in similar but isolated environments. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1837), 20161028.

Pichler, M., Boreux, V., Klein, A., Schleuning, M., & Hartig, F. (2020). Machine learning algorithms to infer trait-matching and predict species interactions in ecological networks. *Methods in Ecology and Evolution*, *11*(2), 281–293.

Pigot, A. L., & Etienne, R. S. (2015). A new dynamic null model for phylogenetic community structure. *Ecology Letters*, *18*(2), 153–163.

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., Kamoun, S., Sumlin, W. D., & Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*(4), 595–609.

Pontarp, M., Brännström, Å., & Petchey, O. L. (2019). Inferring community assembly processes from macroscopic patterns using dynamic eco-evolutionary models and Approximate Bayesian Computation (ABC). *Methods in Ecology and Evolution*, *10*(4), 450–460.

Pontarp, M., & Wiens, J. J. (2017). The origin of species richness patterns along environmental gradients: uniting explanations based on time, diversification rate and carrying capacity. *Journal of Biogeography*, *44*(4), 722–735.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Pudlo, P., Marin, J.-M., Estoup, A., Cornuet, J.-M., Gautier, M., & Robert, C. P. (2015). Reliable ABC model choice via random forests. *Bioinformatics*, *32*(6), 859–866.

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, automatic barcode gap discovery for primary species delimitation. *Molecular Ecology*, *21*(8), 1864–1877.

Quintero, I., & Landis, M. J. (2020). Interdependent phenotypic and biogeographic evolution driven by biotic interactions. *Systematic Biology*, *69*(4), 739–755.

Rabosky, D. L., & Glor, R. E. (2010). Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proceedings of the National Academy of Sciences*, *107*(51), 22178–22183.

Rabosky, D. L., Santini, F., Eastman, J., Smith, S. A., Sidlauskas, B., Chang, J., & Alfaro, M. E. (2013). Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications*, *4*, 1958.

Rangel, T. F. L. V. B., Diniz-Filho, J. A. F., & Colwell, R. K. (2007). Species richness and evolutionary niche dynamics: A spatial pattern–oriented simulation experiment. *The American Naturalist*, *170*(4), 602–616.

Raynal, L., Marin, J.-M., Pudlo, P., Ribatet, M., Robert, C. P., & Estoup, A. (2019). ABC random forests for Bayesian parameter inference. *Bioinformatics*, *35*(10), 1720–1728.

Ree, R. H., Moore, B. R., Webb, C. O., & Donoghue, M. J. (2005). A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution; International Journal of Organic Evolution*, *59*(11), 2299–2311.

Ricklefs, R. E., & Bermingham, E. (2004). Application of Johnson et al.'s speciation threshold model to apparent colonization times of island biotas. *Evolution; International Journal of Organic Evolution*, *58*(8), 1664–1673.

Rosindell, J., Cornell, S. J., Hubbell, S. P., & Etienne, R. S. (2010). Protracted speciation revitalizes the neutral theory of biodiversity. *Ecology Letters*, *13*(6), 716–727.

Rosindell, J., & Harmon, L. J. (2013). A unified model of species immigration, extinction and abundance on islands. *Journal of Biogeography*, *40*(6), 1107–1118.

Rosindell, J., Harmon, L. J., & Etienne, R. S. (2015). Unifying ecology and macroevolution with individual-based theory. *Ecology Letters*, *18*(5), 472–482.

Rosindell, J., Hubbell, S. P., & Etienne, R. S. (2011). The unified neutral theory of biodiversity and biogeography at age ten. *Trends in Ecology & Evolution*, *26*(7), 340–348.

Rosindell, J., & Phillimore, A. B. (2011). A unified model of island biogeography sheds light on the zone of radiation. *Ecology Letters*, *14*(6), 552–560.

Roux, C., Fraïsse, C., Romiguier, J., Anciaux, Y., Galtier, N., & Bierne, N. (2016). Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biology*, *14*(12), e2000234.

Rudman, S. M., Barbour, M. A., Csilléry, K., Gienapp, P., Guillaume, F., Hairston, N. G., Jr., Hendry, A. P., Lasky, J. R., Rafajlović, M., Räsänen, K., Schmidt, P. S., Seehausen, O., Therkildsen, N. O., Turcotte, M. M., & Levine, J. M. (2018). What genomic data can reveal about eco-evolutionary dynamics. *Nature Ecology & Evolution*, *2*(1), 9–15.

Salces-Castellano, A., Patiño, J., Alvarez, N., Andújar, C., Arribas, P., Braojos-Ruiz, J. J., del Arco-Aguilar, M., García-Olivares, V., Karger, D.

N., López, H., & Manolopoulou, I. (2020). Climate drives community-wide divergence within species over a limited spatial scale: Evidence from an oceanic island. *Ecology Letters*, *23*(2), 305–315.

Sanchez, T., Cury, J., Charpiat, G., & Jay, F. (2020). Deep learning for population size history inference: Design, comparison and combination with approximate Bayesian computation. *Molecular Ecology Resources*, *21*(8), 2645–2660.

Sanmartín, I., van der Mark, P., & Ronquist, F. (2008). Inferring dispersal: A Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. *Journal of Biogeography*, *35*(3), 428–449.

Satler, J. D., & Carstens, B. C. (2017). Do ecological communities disperse across biogeographic barriers as a unit? *Molecular Ecology*, *26*(13), 3533–3545.

Schloissnig, S., Arumugam, M., Sunagawa, S., Mitreva, M., Tap, J., Zhu, A., Waller, A., Mende, D. R., Kultima, J. R., Martin, J., Kota, K., Sunyaev, S. R., Weinstock, G. M., & Bork, P. (2013). Genomic variation landscape of the human gut microbiome. *Nature*, *493*(7430), 45–50.

Schrider, D. R., & Kern, A. D. (2018). Supervised machine learning for population genetics: A new paradigm. *Trends in Genetics: TIG*, *34*(4), 301–312.

Sheehan, S., & Song, Y. S. (2016). Deep learning for population genetic inference. *PLoS Computational Biology*, *12*(3), e1004845.

Silva, N. M., Rio, J., Kreutzer, S., Papageorgopoulou, C., & Currat, M. (2018). Bayesian estimation of partial population continuity using ancient DNA and spatially explicit simulations. *Evolutionary Applications*, *11*(9), 1642–1655.

Slater, G. J., Harmon, L. J., Wegmann, D., Joyce, P., Revell, L. J., & Alfaro, M. E. (2012). Fitting models of continuous trait evolution to incompletely sampled comparative data using approximate Bayesian computation. *Evolution; International Journal of Organic Evolution*, *66*(3), 752–762.

Smith, B. T., McCormack, J. E., Cuervo, A. M., Hickerson, M. J., Aleixo, A., Cadena, C. D., Pérez-Emán, J., Burney, C. W., Xie, X., Harvey, M. G., Faircloth, B. C., Glenn, T. C., Derryberry, E. P., Prejean, J., Fields, S., & Brumfield, R. T. (2014). The drivers of tropical speciation. *Nature*, *515*(7527), 406–409.

Smith, M. L., Ruffley, M., Espíndola, A., Tank, D. C., Sullivan, J., & Carstens, B. C. (2017). Demographic model selection using random forests and the site frequency spectrum. *Molecular Ecology*, *26*(17), 4562–4573.

Stone, G. N., Lohse, K., Nicholls, J. A., Fuentes-Utrilla, P., Sinclair, F., Schönrogge, K., Csóka, G., Melika, G., Nieves-Aldrey, J.-L., Pujade-Villar, J., Tavakoli, M., Askew, R. R., & Hickerson, M. J. (2012). Reconstructing community assembly in time and space reveals enemy escape in a Western Palearctic insect community. *Current Biology: CB*, *22*(6), 532–537.

Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, *21*(8), 1789–1793.

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics*, *105*(2), 437–460.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*(3), 585–595.

Tao, Q., Tamura, K., Battistuzzi, F. U., & Kumar, S. (2019). A machine learning method for detecting autocorrelation of evolutionary rates in large phylogenies. *Molecular Biology and Evolution*, *36*(4), 811–824.

Thébault, E., & Fontaine, C. (2010). Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, *329*(5993), 853–856.

Thom, G., Xue, A. T., Sawakuchi, A. O., Ribas, C. C., Hickerson, M. J., Aleixo, A., & Miyaki, C. (2020). Quaternary climate changes as speciation drivers in the Amazon floodplains. *Science Advances*, *6*(11), eaax4718.

Thomas, M. K., Fontana, S., Reyes, M., Kehoe, M., & Pomati, F. (2018). The predictability of a lake phytoplankton community, over time-scales of hours to years. *Ecology Letters*, *21*(5), 619–628.

Thompson, P. L., Guzman, L. M., De Meester, L., Horváth, Z., Ptacnik, R., Vanschoenwinkel, B., Viana, D. S., & Chase, J. M. (2020). A process-based metacommunity framework linking local and regional scale community ecology. *Ecology Letters*, *23*(9), 1314–1329.

Tilman, D. (2004). Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(30), 10854–10861.

Tsuji, S., Maruyama, A., Miya, M., Ushio, M., Sato, H., Minamoto, T., & Yamanaka, H. (2020). Environmental DNA analysis shows high potential as a tool for estimating intraspecific genetic diversity in a wild fish population. *Molecular Ecology Resources*, *20*(5), 1248–1258.

Turon, X., Antich, A., Palacín, C., Praebel, K., & Wangensteen, O. S. (2020). From metabarcoding to metaphylogeography: Separating the wheat from the chaff. *Ecological Applications*, *30*(2), e02036.

Valente, L., Phillimore, A. B., Melo, M., Warren, B. H., Clegg, S. M., Havenstein, K., Tiedemann, R., Illera, J. C., Thébaud, C., Aschenbach, T., & Etienne, R. S. (2020). A simple dynamic model explains the diversity of island birds worldwide. *Nature*, *579*(7797), 92–96.

Valente, L. M., Phillimore, A. B., & Etienne, R. S. (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galápagos islands. *Ecology Letters*, *18*(8), 844–852.

Varin, C., Reid, N., & Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica*, *21*(1), 5–42.

Vellend, M. (2005). Species diversity and genetic diversity: Parallel processes and correlated patterns. *The American Naturalist*, *166*(2), 199–215.

Vellend, M. (2016). *The theory of ecological communities (MPB-57)*. Princeton University Press.

Wang, S., Chen, A., Fang, J., & Pacala, S. W. (2013). Why abundant tropical tree species are phylogenetically old. *Proceedings of the National Academy of Sciences*, *110*(40), 16039–16043.

Warren, B. H., Simberloff, D., Ricklefs, R. E., Aguilée, R., Condamine, F. L., Gravel, D., Morlon, H., Mouquet, N., Rosindell, J., Casquet, J., Conti, E., Cornuault, J., Fernández-Palacios, J. M., Hengl, T., Norder, S. J., Rijsdijk, K. F., Sanmartín, I., Strasberg, D., Triantis, K. A., … Thébaud, C. (2015). Islands as model systems in ecology and evolution: Prospects fifty years after MacArthur-Wilson. *Ecology Letters*, *18*(2), 200–217.

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, *7*(2), 256–276.

Weinstein, B. G., Tinoco, B., Parra, J. L., Brown, L. M., McGuire, J. A., Stiles, F. G., & Graham, C. H. (2014). Taxonomic, phylogenetic, and trait Beta diversity in South American hummingbirds. *The American Naturalist*, *184*(2), 211–224.

Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, *16*(2), 97–159.

Wright, S. (1938). The distribution of gene frequencies under irreversible mutation. *Proceedings of the National Academy of Sciences of the United States of America*, *24*(7), 253–259.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114.

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*, *19*(3), 395–420.

Xu, L., van Doorn, S., Hildenbrandt, H., & Etienne, R. S. (2020). Inferring the effect of species interactions on trait evolution. *Systematic Biology*, *70*(3), 463–479.

Xu, W., Liu, L., He, T., Cao, M., Sha, L., Hu, Y., Li, Q., & Li, J. (2016). Soil properties drive a negative correlation between species diversity and genetic diversity in a tropical seasonal rainforest. *Scientific Reports*, *6*, 20652.

Xue, A. T., & Hickerson, M. J. (2020). Comparative phylogeographic inference with genome-wide data from aggregated population-pairs. *Evolution*, *74*(5), 808–830.

BIOSKETCH

Members of the **iBioGen consortium** and collaborators promote methodological unification and theoretical synthesis for studying island biodiversity using community-scale DNA sequence data, such as can be obtained from contemporary metabarcoding studies. To this end, we develop and apply novel methods that unify across scales of biological organization for inferring ecological and evolutionary processes from multidimensional biodiversity data that include a genetic component. Further information about the iBioGen project and its partners may be found on the project website: https://www.ibiogen.eu/.

How to cite this article: Overcast, I., Achaz, G., Aguilée, R., Andújar, C., Arribas, P., Creedy, T. J., Economo, E. P., Etienne, R. S., Gillespie, R., Jacquet, C., Jay, F., Kennedy, S., Krehenwinkel, H., Lambert, A., Meramveliotakis, E., Noguerales, V., Perez-Lamarque, B., Roderick, G., Rogers, H. … Morlon, H. (2023). Towards a genetic theory of island biogeography: Inferring processes from multidimensional community-scale data. *Global Ecology and Biogeography*, *32*, 4–23. https://doi.org/10.1111/geb.13604