Process-explicit models reveal the structure and dynamics of biodiversity patterns

Julia A. Pilowsky1, Robert K. Colwell2,3,4,5, Carsten Rahbek2,6,7,8, Damien A. Fordham1,2,6 *

With ever-growing data availability and computational power at our disposal, we now have the capacity to use process-explicit models more widely to reveal the ecological and evolutionary mechanisms responsible for spatial-temporal patterns of biodiversity. Most research questions focused on the distribution of diversity cannot be answered experimentally, because many important environmental drivers and biological constraints operate at large spatiotemporal scales. However, we can encode proposed mechanisms into models, observe the patterns they produce in virtual environments, and validate these patterns against real-world data or theoretical expectations. This approach can advance understanding of generalizable mechanisms responsible for the distributions of organisms, communities, and ecosystems in space and time, advancing basic and applied science. We review recent developments in process-explicit models and how they have improved knowledge of the distribution and dynamics of life on Earth, enabling biodiversity to be better understood and managed through a deeper recognition of the processes that shape genetic, species, and ecosystem diversity.

INTRODUCTION

The patterns of biodiversity we observe at different temporal and spatial scales result from the key evolutionary and ecological processes of speciation, ecological interaction, adaptation, movement, and extinction, acting separately or in concert (1). These processes can be stochastic or forced by natural drivers of environmental change (e.g., plate tectonics and paleoclimate change) or by human drivers, such as invasive species, land use, pollution, and harvesting (2). However, the interplay among these processes and their drivers is complex (3), and different sets of circumstances can produce similar patterns. This ambiguity has made it difficult to discern which ecological and evolutionary processes and drivers have shaped current-day patterns of biodiversity based on empirical data alone (4). Fortunately, key advances in process-explicit models over the past 50 years are now enabling the processes and drivers responsible for contemporary patterns of biodiversity to be disentangled in space and time. Here, we show how these advances in biodiversity modeling are revealing the generalizable mechanisms responsible for the distributions, abundances, and diversity of life on Earth and how they are strengthening basic and applied science, resulting in improved guidelines for the management of nature.

Process-explicit models in ecology and evolution represent the dynamics of a biological system as explicit functions of the events that drive change in that system (5). By causally linking current patterns to the past events that produced them (Fig. 1), process-explicit models help achieve a deeper understanding of the chain of causality leading to current-day spatial patterns of biodiversity, including human diversity (6). These models allow contested ecological and evolutionary theories to be assessed, enabling biodiversity to be understood and managed more effectively through a deeper recognition of the processes of genetic-, species-, and ecosystem-level endangerment and collapse (7).

Models that are process-explicit provide platforms for directly integrating ecological and evolutionary theory into conservation and environmental science (8), enhancing knowledge of the effects of biodiversity and its drivers on the functioning of species and ecosystems (9), and strengthening projections of biodiversity in a changing world (10), resulting in improvements to conservation management and policy (11). For example, process-explicit models derived from the neutral theory of biodiversity (12) were some of the first models to show that rare species are less frequent in island communities than in adjacent mainland communities (13), providing important new information to conservation policy-makers regarding vulnerability to human-driven environmental change (14). Process-explicit models of the neutral theory of molecular evolution, which simulate rates of genetic drift as products of effective population size and generation length (15), enabled conservation geneticists to study the behavior of neutral alleles to better understand why extinction risk increases for species with small population sizes (16). A stronger integration of ecological and evolutionary theory in conservation science using process-explicit modeling promises to further link the evolution of species traits at the individual level to the dynamics of communities and the overall functioning of ecosystems (17). Together, these advances are improving knowledge of how climatic and environmental changes have shaped species assemblages in the past, strengthening confidence in projections of biodiversity’s future (7).

Recent reviews have established important benefits of process-explicit modeling approaches in macroecology (5, 18), ecosystem ecology (17), conservation science (10), and related disciplines. These studies highlight a need to use process-explicit models for managing ecosystems (17), improving theory (5, 18), and predicting species’ range shifts under ongoing and future climate change (10). However, there has been no synthesis of the broader uses of process-explicit models for unraveling the biological mechanisms responsible for shaping patterns of biodiversity in space and time in response to...
In the 1980s, development of coalescent models of simulated genealogies (28) enabled the diversification of lineages to be studied in space and time (29), giving rise to the field of phylogeography. These early models showed how lineages can diverge without geographic isolation, illustrating potential mechanisms of sympatric speciation. More recently, they have been used to show how pathogens can rapidly evolve as they spread through a network of hosts (30), enriching fundamental understanding of past, current, and future disease dynamics (31). The latest generation of coalescent models can reconstruct genomic erosion in endangered species (32) and rapid directional selection (33) in response to subcentennial periods of environmental change.

The 1990s saw the advent of dynamic global vegetation models (DGVMs): process-explicit models that replicate global patterns of vegetation by simulating the growth and mortality of plant functional groups under different climatic conditions (34). This development enabled predictions of the capacity of the biosphere to store carbon (35) and produce crops (36) under current and future climate conditions. Today, DGVMs are being used to inform regional-to-global policies on food security, greenhouse gas emission scenarios, and the maintenance of ecosystem services (37). They can account for the effects of herbivory and fire regimes on vegetation structure (38), allowing the impacts of competing land management strategies to be compared (11, 39).

In the early 2000s, models that integrate the evolutionary processes of speciation and adaptation with the ecological processes of movement, extinction, and interaction began to be developed. By providing a mechanistic understanding of the physical and biological processes that shape Earth’s biodiversity, these models have aimed to illuminate the origins of biodiversity through direct tests of competing scenarios (40). Many of these theories, established long ago by early naturalists (41–43), could not be directly tested with simpler process-explicit models or phenomenological approaches. Today, eco-evolutionary simulators provide opportunities to achieve new levels of realism in projections of assemblage dynamics under past and future global change (44).

The most recent developments in process-explicit modeling, which simulate multiple processes and patterns of biodiversity using complex mathematical components and logical algorithms, have resulted from a rapid rise in computational power following the turn of the 21st century (44–46). This advance, coupled with wider access to large ecological, genomics, and satellite-based remote sensing datasets, has enabled the generation and increasingly frequent application of a broad variety of process-explicit models in ecological and evolutionary studies, parameterized or validated with more data and based on more-realistic assumptions than previously possible. Despite this accelerated expansion, the development and application of process-explicit models have followed an opportunistic path, with little strategy or coordination (47).

To address this current shortfall, we provide here a much-needed review of recent developments in process-explicit models, outlining considerations for researchers who contemplate building process-explicit models to evaluate the mechanisms that govern the structure and dynamics of past, present, and future biodiversity. We scrutinize the processes of codifying theory into models, identify key scientific advances from simulation outputs, and illustrate with examples how process-explicit models can safeguard future biodiversity.

**PROCESS VERSUS PATTERN**

Narrative accounts (42), correlative studies (48), and experiments (49) lead to hypotheses about the underlying causes of biodiversity change, and theoretical models demonstrate possible mechanisms (50). In comparison, spatiotemporal process-explicit models can directly assess and disentangle competing theories for drivers of biodiversity.
biodiversity, helping to elucidate interactions among underlying ecological and evolutionary mechanisms and drivers (5). An example of competing theories for biodiversity dynamics and resultant patterns is the contrast between niche (51) and neutral theory (12). The former focuses on the role of environmental determinism, while the latter focuses on contingent and stochastic determinants of biodiversity dynamics.

Process-explicit models differ from pattern-based models by generating predictions based on explicit causal relationships between environmental drivers and ecological and evolutionary responses, rather than inferring implicit causal relationships based on correlations between observed and modeled patterns (52). A physiological model, for example, is process-explicit if it characterizes the occurrence of a tree species in a landscape based on where the tree can minimize water stress. In contrast, a model is phenomenological (or correlative) if it maps the tree’s occurrence based on the statistical relationship between annual precipitation and observations of occurrence, because no processes linking precipitation and fitness are specified. The process-explicit model allows patterns (e.g., a contraction of the tree species’ range) to be connected to processes that cause them (e.g., an increase in seedling mortality in a drought), while the phenomenological model cannot explicitly link a changing pattern to a causative agent (53). Similarly, a phenomenological model that hindcasts plant functional types on the landscape based on correlations between climate and pollen records cannot link pattern and process in the same way as a DGVM that hindcasts plant nutrient cycling and competition over the same period (54).

Phenomenological models and experimental observations sometimes find strong or unexpected correlations that can suggest the mechanisms that produce them. Proposed mechanisms can be used to build process-explicit models that can then be tested against observed patterns (7). Studies of the effect of biodiversity on ecosystem function offer an example of this ontology (Fig. 2). Effects of depauperate plant richness on ecosystem function were first observed empirically in experimental chambers and plots, which led to the proposed mechanism of niche complementarity, which, in turn, became the basis for mechanistic models of ecosystem function (55). In this way, phenomenological and experimental analysis can provide important insights into the workings of nature that can be tested using process-explicit models.

**REVEALING STRUCTURE AND DYNAMICS**

Process-explicit models can operate at diverse levels of biological organization, ranging from the gene to the ecosystem (Fig. 3). The level of biological organization that is simulated—genetic, species, or ecosystem diversity—has, to date, dictated the number and combination of possible biotic processes that are modeled (7). The five primary processes responsible for the origin, structure, and dynamics of biodiversity are speciation, ecological interaction, adaptation, movement, and extinction (including population extirpation). In this context, ecological interactions encompass both interspecific species interactions (competition, predation, herbivory, parasitism, and mutualism) and ecosystem processes (nutrient cycling, photosynthesis, stability, etc.).

Ecosystem- and population-level models were the earliest process-explicit models. They generally include ecological interaction and local- to range-wide extinction processes (Fig. 3), but not movement, speciation, or adaptation. In contrast, more recently developed community-level models simulate all five primary biotic processes (46). These and individual-based models are becoming more frequently used to unravel biological mechanisms that underpin spatiotemporal patterns of biodiversity (Fig. 3). These advances promise to lead to a greater awareness of the importance of evolutionary processes in shaping biodiversity (52).

**Genetic diversity**

Although coalescent models have simulated genetic diversity—trait inheritance within species—for 40 years (28), early approaches did not model differences in DNA among individuals in space and time. This advance was not made until the beginning of the 21st century (Fig. 3) with the advent of a spatially explicit simulation framework.
for population genetics: the serial-genetic simulator SPLATCHE (SPatial and Temporal Coalescent in a Heterogeneous Environment). The first studies to use SPLATCHE found that range expansions in heterogeneous environments produce genetic diversity patterns contingent on the geographical origin of the expansion, allowing spatially explicit genetic models to trace back the origin points of range expansions (56). Subsequently, coalescent-based process-explicit models have been frequently used to infer the effects of species’ range expansions, contractions, and shifts on patterns of genetic diversity, using ancient and modern DNA. They have revealed that genetic diversity declines toward the leading edge of a species range more steeply than predicted by neutral theory (57) and that rapid range contractions conserve more genetic diversity in refugial populations than slow range contractions (58). These models have also shown that present-day isolation of a population is a poor indicator of the past diversity of the lineage and historical barriers to gene flow (59) and that rapid warming events can reconfigure species assemblages (60). Together, these reconstructions of past patterns of genetic diversity using process-explicit models are helping to improve projections of future patterns of genetic diversity by parameterizing known responses to environmental shifts (61).

Virtual genomes can be simulated to test and refine theories of genetic diversity. These genomes are simulated with mutation, migration, and divergence on computer-generated landscapes using a priori mutation rates and dispersal patterns. This approach has been used to simulate species’ range expansions, revealing that introgression [transfer of genetic information from one species to another as a result of hybridization (62)] is likely to occur from the resident population to the invading population, regardless of the relative densities of the resident and invader populations (63). Simulations of virtual genomes have also shown that new mutations near the leading edge of an expanding range have a higher frequency and wider spatial distribution than in a stationary population (64). This result suggests that spatially expanding populations have an increased rate of evolution at their frontier (64), with important implications for the management of invasive populations and range-shifting native species.

**Species diversity**

Process-explicit models can be used to decipher patterns of species diversity at the level of the individual, species, or community, and findings underpinning the operation of biological processes at these different levels of species diversity can reinforce or amplify one another. For example, an individual-level model can elucidate the evolution of optimal dispersal strategies within a single habitat island (65), a population-level model can reveal species diversity patterns across a chain of islands shaped by different dispersal strategies (66), and a community-level model can infer dispersal strategies in different functional groups, based on diversity across an entire region (67). In this way, process-explicit models at these three levels of organization allow us to investigate and potentially to integrate the impact of movement on species diversity patterns at multiple biological scales.

Processes can be modeled at the level of the individual organism with agent-based models (68) and physiological approaches (69). The former can potentially capture any of the five fundamental biotic processes responsible for biodiversity and can generate complex population- and community-level phenomena that arise from ecological interactions among individuals (70). For example, individual-based models of initial colonization in a range expansion or shift have shown that the interaction of local adaptation with timing (71) and speed (72) of colonization can alter the expected distribution of a species along an environmental gradient.

However, models built at the individual level can be computationally intensive, particularly if they simulate complex eco-evolutionary processes for many populations of individuals. Moreover, they can be difficult to parameterize and validate (Fig. 4), because data on biotic processes like movement and other attributes are often unavailable at the level of the individual. The computational demands of these models have led some researchers to use machine learning techniques (as emulators) to generalize process-explicit model behavior post hoc at small scales and apply those generalizations to larger scales (73). Others have used virtual landscapes to simulate and explore the population-level effects of different movement strategies, requiring neither biotic nor environmental data for parameterization (74). This approach, which allows for the simulation of data-poor processes at the individual level, has shown that range shifts can be accelerated by the evolution of greater dispersal ability in marginal habitats (65).

Physiological models, such as NicheMapper (74), and forest gap models, such as ForClim (75) and PHENOFIT (76), simulate only local- to range-wide extinction in animals and trees, respectively, making them computationally less intensive than individual-level models at large spatial scales. These approaches assume that if environmental conditions are suitable given an organism’s physiological traits, it will persist; otherwise, it will die. These models are built on physiology and thermal tolerances, which are used to predict where individuals can survive. Physiological models can refine projections from phenomenological models of species distribution by identifying locales where heat stress will cause local extinction, informing conservation management (53, 77).

Individual-level models are, nevertheless, often constrained to ecological and evolutionary processes at local extents, often failing to account for potentially important coarser-scale processes that...
can affect species diversity. Population models, which find their roots in simple logistic growth equations or matrix population models (78), can simulate movement and mortality in a network of populations extending across a species range (79, 80). They can simulate trait values and genes, thus incorporating adaptation or speciation among populations over time, uncovering interactive effects of adaptation and dispersal on distributions of phenotypes (81). Although these models usually feature one or a few focal species, they can be used to simulate many populations of interacting species, capturing ecological interactions and community dynamics (66). For example, a model of competing and evolving populations has shown that certain syndromes of life history traits (mating system and dispersal ability) outcompete others—a mechanistic prediction that fits with empirical observations in plants (81).

Pathways to extinction are difficult to detect and disentangle phenomenologically (82) because they are complex, often starting long before the extinction event, resulting from biological responses to natural and human-induced factors that operate at multiple spatiotemporal scales (80). Linking population models to correlative species distribution models to address well-recognized limitations of pattern-based approaches (83) is allowing the processes of movement, extinction, and—most recently—adaptation to be simulated over multiple millennia (84). This approach is revealing how ecological strategies, and demographic and evolutionary traits, interact dynamically with past environmental change and human-driven factors to cause the decline and eventual extinction of species (80).

Biodiversity loss can be modeled for groups of interacting species using community-level models. Process-explicit models at the community level simulate biogeographical dynamics with species as functional units within the simulation (52). Unlike population models, which typically have species or population distributions as their outputs, or ecosystem models, which generally produce maps of ecosystem function or plant functional guilds (see below), these community-level biogeographical models usually generate species richness maps and range size frequencies (46).

Most community-level process-explicit models encompass all of the five biotic processes that drive biodiversity, making them aptly suited for testing differing hypotheses about the underlying causes of patterns of biodiversity, including how lineages diversify over space and time. For example, community-level process-explicit models have been used to determine whether neutral theory can explain empirical patterns of reef community dynamics, finding support for the theoretical expectation that range size should increase with dispersal ability (85). However, models of community-level processes not only are used to answer theoretical questions about biodiversity but also can be applied directly to real-world ecological systems to understand patterns of species richness (86), community assembly (87), and diversity loss (88) in a changing world. Diversification models with simple parameterization have applications in conservation biology, including identifying the effects of environmental change on biodiversity hotspots (89) and predicting the loss of species in a community after habitat destruction (88).
their complexity, these process-explicit models of biogeographical dynamics can be validated (Fig. 4) using targets of current-day range size frequency distributions (45).

**Ecosystem diversity**

Ecosystem diversity models simulate the structure of functional groups of terrestrial and marine organisms. The coexistence and interactions of these groups are used to map the distribution of ecosystems (90). Interactions among terrestrial autotrophs and the abiotic environment are modeled with DGVMs (34), while fisheries management models (91) and general ecosystem models (92) also include primary and secondary consumer dynamics, enabling simulation of energy transfer through food webs. These ecosystem diversity models are being used to forecast and manage ecosystem services, including carbon storage (93), clean water supply (94), and food security (95) in a changing world. They have shown that freshwater supply will be reduced under future warming to the detriment of terrestrial ecosystem functioning (96), that increased hurricane frequency threatens the structure and productivity of reef-fish communities (97), and that habitat fragmentation affects the trophic structure of ecosystems (98). Furthermore, process-explicit ecosystem models have shown that forest function is more resilient to warming events in high- than in low-diversity forests (99), illustrating causative mechanisms for experimental observations (55).

DGVMs simulate the distribution of plant functional types as well as their fluxes of carbon, water, and nutrients through the environment (34), enabling them to simulate dynamic feedbacks between the biosphere and the climate when coupled to climate models (100). This coupling of models has uncovered important interactions between climate, CO2, and ecosystem function, including evidence that a positive interaction between plant productivity and elevated levels of CO2 can potentially offset the negative effects that climate change and, more specifically, increased aridification can have on productivity (101). Moreover, by hindcasting ecosystem diversity dynamics over glacial-interglacial cycles, DGVMs have disentangled many of the effects of climate on ecosystem structure (102). For example, modeling the interaction between deglaciating and megaherbivore die-off following the last glacial maximum reveals how high-latitude mammoth steppe—Earth’s most extensive biome at the time—was converted to a taiga-tundra ecotone (38).

While more complex general ecosystem models can simulate the entire ecosystem, from phyto- and zooplankton to apex carnivores, capturing complex food web dynamics, they do not as yet include two-way interactions with climate (92, 98). Consequently, they are frequently used to test theories regarding ecosystem structure, including relationships between heterotrophic biomass and net primary productivity (92), and to determine the impact of recent land-use change on ecosystem function (98). The application of these ecosystem-level models in fisheries management has uncovered crucial ecosystem services provided by coral reefs, including calcium carbonate deposition and coastal protection, showing how overfishing disrupts these services to nature and people (103).

**RELATIONSHIP TO DATA AND THEORY**

Process-explicit models have a variety of relationships with data and theory (fig. S1). Some process-explicit models are theory driven: Their purpose is to explore the implications or applications of an ecological theory, such as the neutral theory of biodiversity (12), the species-area relationship (104), or the general dynamic theory of island biogeography (105). Others are theory scaffolded: Their purpose is to understand an ecological system empirically and to use theory as a scaffold by which to structure the model and interpret its outputs (101).

While process-explicit models are diverse in structure (60, 84), they exist on two distinct continua, based on (i) their use of empirical data for parameterization and (ii) how they are verified and/or validated (Fig. 4). Empirical data are not necessary to build and run a process-explicit model. Indeed, many theory-driven models use arbitrary values for parameters and explore the interactions and patterns that result from the model (87). These models are at one end of a parameterization continuum. Further along the continuum are models that use either biotic data (such as genetic sequences or species occurrence) or environmental data (such as spatiotemporal climatic fluctuations or bathymetry change) to parameterize models, but not both (46). The next category of models includes those that use biotic and static environmental data (106), followed by models that use biotic and dynamic environmental data (34). In the last two cases, biotic data represent a single level of biological organization: gene, individual, population, community, or ecosystem (7). At the most extreme end of the parameterization continuum lie models that use dynamic environmental data and biotic data to simulate processes across multiple levels of biological organization: for example, simulating individual-level movement (based on seed dispersal by wind) and population-level mortality (based on survival across individuals) (107).

A second distinct gradient specifies how data are used for verification and validation in process-explicit models (Fig. 4). Verification is a check to ensure that the implemented model meets the primary theoretical assumptions it has been built to represent. In contrast, validation evaluates the level of correspondence between the implemented model and the study system (108). At one end of the verification and validation continuum, model outputs are not verified or validated at all. Moving up the continuum, output patterns can be verified for congruence with theory by comparing model outputs with well-established theoretical relationships, such as the mid-domain effect (109). Models can be validated through visual inspection of patterns based on observational data, using nonstatistical procedures (110). Statistical validation allows model outputs to be evaluated with patterns of empirical data, by means of measures such as coefficient of determination ($r^2$), root mean square error, or true skill statistic (102). At the most data-heavy end of the verification and validation continuum lies multivariate statistical validation (59), in which models are evaluated on the basis of their ability to simultaneously demonstrate goodness of fit to multiple empirical patterns. This demanding level of validation is now being applied to pattern-oriented modeling (an emerging and powerful technique in data science), in which mechanisms governing the structure and dynamics of biodiversity are identified by converging model simulations to independent multivariate validation targets (45, 80).

Figure 4 shows how process-explicit models with diverse relationships to data can be used to decipher the mechanisms underlying the structure of biodiversity. Models that use little or no empirical data can be used to test ecological and evolutionary theories, such as modes of speciation (111). These primarily theory-driven models are useful even when biological data are not available for validation; for example, data-free process-explicit models can test the sensitivity of model outputs to underlying processes (111), distinguishing...
metapopulation dynamics from neutral dynamics or random community assemblage (112). Theory-scaffolded models with complex parameterization often have greater explanatory power, particularly if they use more than one level of biotic data for parameterization and validation, and if they simulate dynamic drivers of global change affecting the spatial structure of biodiversity. These additional data inputs can allow otherwise necessary model assumptions to be relaxed, such as an assumption of unlimited movement (107) or static human land use (39), while multivariate validation targets (despite being, so far, rarely used) provide more stringent tests of model simulations.

SAFEGUARDING BIODIVERSITY

Sustainable management of biodiversity has been recognized as a policy goal for 30 years (113); however, progress in halting the decline and degradation of biodiversity has been limited (114). Reasons for failing to reduce biodiversity loss are complex, reflecting long-lasting knowledge gaps on biodiversity dynamics (47), as well as insufficient integration of biodiversity science in policy making (115) and lack of motivation to deliver the required biodiversity changes (116). An incomplete understanding of the mechanisms that govern the structure and dynamics of biodiversity and a tendency to use correlative rather than process-explicit approaches to forecast the future of biodiversity in a changing world (10) have constrained capabilities to set productive biodiversity targets, develop cross-cutting solutions for restoring nature, and obtain national commitments to biodiversity conservation.

Process-explicit models have a diverse range of applications, including formulating and assessing potential solutions for mitigating future genetic-, species-, and ecosystem-level collapse. Currently, for example, the paleorecord is being used to identify biological mechanisms that mediate responses to climate- and human-driven change using process-explicit models (117). These paleo-models can disentangle past determinants of genetic diversity, range shifts, species richness, and ecosystem structure and function. By specifying the causal processes that underpin biodiversity change, they can provide the context needed to improve confidence in predictions of biodiversity’s future (7), leading to improved computational platforms for setting biodiversity targets and better solutions for mitigating adverse changes to biodiversity (8).

The genetic signatures of demographic responses of species to environmental changes can be decoded using genetic simulation models (6) to better manage future biodiversity (118). For example, process-explicit models of gene flow, which allow demographic trends and gene flow to be reconstructed (16), are establishing the importance of intraspecific genetic diversity for resilience to accelerated climatic change (119). There is now a push to use this technique more widely to improve knowledge of how rapid climatic change affects patterns of genetic diversity (61). In the absence of ample genetic samples, process-explicit models can still be used to test theories central to conservation genetics using virtual genetic sequences and landscapes (120) to deliver valuable information for conserving future genetic diversity (121).

Historical context is crucial for understanding the threat of future declines in species distributions. Process-explicit models constructed at the individual and population level can be used to identify demographic processes that cause range shifts for a species or suites of species in response to climatic and environmental drivers, improving species threat assessments (77). Because individual-based models often operate at a level of detail that is not necessary for simulating range dynamics across large extents, process-explicit population-level models are more commonly used to project past and future range dynamics. These population-level models can be used to identify ecological traits that cause species to be differentially prone to regional and range-wide extinction (122) and to evaluate the efficacy of current methods for identifying threatened species (123). Population-level models that incorporate adaptation as a process have been influential and instructive in revealing the role of gene flow along ecological selection gradients, and its inhibiting effect on local adaptation to environmental change (124).

Hotspots of biodiversity are of particular conservation concern because they support high concentrations of species, particularly endemics (125). Process-explicit models built and validated at the community level to simulate geographical patterns of species richness and endemism can identify mechanisms central to the maintenance of past and contemporary hotspots of species richness (126), providing a framework for assessing vulnerability to future climate and environmental change (Fig. 5). If simulations can capture community-level responses to realistic temps and magnitudes of future global change, these new predictive approaches will benefit 21st century environmental management and conservation (7).

To illustrate the state of the art in broad-scale modeling of biodiversity and its potential application for biodiversity conservation, we offer an example of a community-level, process-explicit model that incorporates all five biological processes that govern the structure and dynamics of biodiversity in a temporally dynamic environment (46). The model was designed to simulate the geographic

Fig. 5. Models for predicting continental species richness. Community-level biogeographical models (46), driven by interactions between climate and biological processes, can incorporate all five biological processes that govern biodiversity: movement, extinction, ecological interaction, adaptation, and speciation. Model outputs can simulate maps of current-day and future species richness and endemism (rarity-weighted species richness). Top plot shows temperature across thousands of years (ka). Image of finches adapted from Charles Darwin.
distributions and patterns of overlap of species ranges in response to the past 800,000 years of climate change in South America (Fig. 5). In this model, evolutionary niche dynamics drive range expansion and fragmentation (leading to speciation), adaptation to climatic conditions, and extinction. Combinations of parameter settings (dispersal distance, evolutionary rate, time for speciation, and intensity of competition) for virtual species were chosen to priori, producing many different potentially plausible range maps. Although not directed by any empirical validation targets, the emerging maps closely resembled contemporary species richness of major South American taxa. Combinations of parameters that closely reproduced the current-day biodiversity of South American avifauna (including hotspots of species richness and endemism) showed that low rates of adaptation to past climatic change were required to reconstruct observed patterns of species richness. In the future, such community-level simulation models (built to simulate the past and validated in the present; Fig. 5) could be parameterized with climate forecasts to predict strongholds of species richness under future climates. The subsequent results could be used to guide the protection and future management of biodiversity.

By identifying the biological mechanisms, drivers, and their interactions that mediate changes in ecosystem structure and function, process-explicit models can help safeguard the services ecosystems provide to nature and people. Early ecosystem models were used to investigate the effects of increased atmospheric carbon on vegetation communities (96). More recent models have incorporated complex interactions between multiple drivers of global change and ecosystem-level processes, including the effects of agriculture and land-use change (93). This research has strengthened knowledge of the drivers and responses that underpin change in ecosystem structure and function (36, 93), improving projections (127) and informing protocols for assessing ecosystem threat status (128). For example, DGVMs have shown mechanistically how 20th century agriculture caused a 24% reduction in global vegetation and a 10% reduction in global soil carbon (93). A better understanding of processes of ecosystem change enables the simulation of the effects of current and future climatic and environmental change (including altered fire regimes) on important ecosystem services, such as agricultural productivity, freshwater availability, and timber production (36, 39).

Climate projections are currently made using models characterized by complex system dynamics, including interactions and feedbacks between the atmosphere, ocean, land, and society (129). While analogous models for projecting biodiversity change have typically been simpler than approaches used in climate science, general ecosystem models (92) and process-based community assemblage models (44) offer new and more robust methods for projecting the future distribution of life on Earth. These next-generation biodiversity models, which explicitly capture the structure and dynamics of biodiversity, will strengthen our capacity to set achievable biodiversity targets that promote engagement and investment where change is needed.

LOOKING FORWARD

Although phenomenological models are a crucial first step toward understanding the potential determinants of current and past spatial patterns of biodiversity, process-explicit models are needed to identify causal processes that govern the structure and dynamics of biodiversity, and to exclude those that do not. Increased open access to curated georeferenced occurrence records, dated fossils, libraries of genetic sequences, and climate simulations will continue to provide innovative opportunities to apply process-explicit models, especially to connect inferences of past responses of biodiversity to different rates and magnitudes of contemporary climate and environmental change (117). These opportunities include testing key assumptions of existing biodiversity models—such as the common assumption that processes driving changes in biodiversity are scale invariant (130)—and competing theories for large-scale biodiversity patterns, including geographical gradients in species richness (44).

Continuous simulations of the transient late Quaternary climate are needed, ideally at fine spatial resolutions, to determine population-, species-, community-, and ecosystem-level responses to abrupt (as well as gradual) climatic change using process-explicit models (131). The TRaCE21ka experiment based on the Community Climate System Model version 3 (132) has bridged this gap, but it spans only the past 21,000 years. Higher spatiotemporal resolution paleoclimate simulations from Earth systems models before 21,000 years ago that include solar flux, ice sheet extent, and sea level changes will provide a more thorough understanding of the mechanisms responsible for spatiotemporal patterns of biodiversity at evolutionary time scales (7). Statistical emulators of climatic change will be useful in filling this data and knowledge gap (133), particularly in the Southern Hemisphere, for which there is a paucity of high-resolution simulated data before the last glacial maximum (134). Including better reconstructions of solar variability, volcanic eruptions, and land use during the Holocene in transient simulations of Earth’s climate will provide a more complete picture of more recent temporal change in regional climates and the biodiversity they support.

Integrating paleoclimatological and neoclimatological perspectives into process-explicit models is key to contextualizing the present and anticipating and visualizing ecological responses to future global change (117). Emerging genomic techniques are allowing genetic diversity and effective population size to be estimated over short periods (<100 years) of environmental change, providing inferences of eco-evolutionary change to recent and/or punctuated disturbance events (32, 33, 135) that can feed directly into process-explicit models of range collapse and population declines. Importantly, projections of recent climate, vegetation, and land-use change have been harmonized with ancient projections, allowing their effects on biodiversity to be characterized continuously in process-explicit models that run from as far back as 21,000 years ago to the present day (136) and, in some cases, into the future (137).

Adaptation was first incorporated into spatial process-explicit models in the early 2000s (138) and has since become more common in ecological and evolutionary models. However, it still remains the most infrequently modeled biological process. A more regular integration of adaptation into process-explicit models of climate change responses will benefit from taxonomically diverse datasets of historic DNA that are readily available today (139) and from technological advances that allow ancient DNA to be used to reconstruct shifts in genetic diversity and adaptations to large-magnitude and abrupt climatic change (140). Adding community dynamics to population models and demography to community models will also strengthen projections of biodiversity change. Metacommunity models with simplified food webs can bridge this gap by modeling demographic interactions between populations of multiple species in a spatiotemporally explicit manner (141). Community-level models can integrate a higher level of biological organization by combining
ecosystem-level drivers such as fire with processes of plant community assembly (142).

Achieving more detailed mechanistic understandings of patterns of biodiversity—from the gene to the ecosystem level—will require a greater focus on rigorous statistical validation of process-explicit models using independent multivariate data that are spatiotemporally explicit. In systems where theory is not yet well developed, empirical data for model parameterization are needed to simulate realistic outputs. However, as the mechanisms underpinning a system’s biodiversity become better understood, model outputs will be simulated using theory alone. Realistic predictions generated from a strong theoretical framework are the pinnacle that ecologists and evolutionary biologists should aim for when wielding process-explicit models.

Process-explicit models have been instrumental in improving knowledge of the distribution of life on Earth, revealing complex causal processes for contemporary patterns of biodiversity that could not be discerned from experimental approaches or phenomenological models. A deeper recognition of the structure and dynamics of organisms, communities, and ecosystems in process-explicit models is helping to protect and restore biodiversity by formulating remedies to existing problems and countering undesirable future changes.

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abj2271

REFERENCES AND NOTES

42. A. von Humboldt, Personal Narrative of Travels to the Equinoctial Regions of the New Continent During the Years 1799–1804 (G. Bell, 1877).


Acknowledgments
Funding: D.A.F. acknowledges funding from the Australian Research Council (FT140101192, DP180102392) and a residency fellowship from Danmarks Nationalbank. C.R. received funding from DNRF-CMEC (DNRF96) and from Villum Fonden (grant no. 25925). Author contributions: J.A.P., D.A.F., and C.R. conceived the idea for the paper. J.A.P. did the literature review. All authors contributed to writing the manuscript. Competing interests: The authors declare that they have no competing interests. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.
Process-explicit models reveal the structure and dynamics of biodiversity patterns
Julia A. PilowskyRobert K. ColwellCarsten RahbekDamien A. Fordham

Sci. Adv., 8 (31), eabj2271. • DOI: 10.1126/sciadv.abj2271

View the article online
https://www.science.org/doi/10.1126/sciadv.abj2271
Permissions
https://www.science.org/help/reprints-and-permissions