# ECOGRAPHY

*Review and synthesis*

WINNER

# Understanding ecological change across large spatial, temporal and taxonomic scales: integrating data and methods in light of theory

## Giovanni Rapacciuolo and Jessica L. Blois

*G. Rapacciuolo (https://orcid.org/0000-0003-1494-9017) ✉ (giorapac@gmail.com) and J. L. Blois, Life and Environmental Sciences, Univ. of California Merced, 5200 N. Lake Road, Merced, CA 95343, USA. GR also at: California Academy of Sciences, San Francisco, CA, USA.*

The difficulty of integrating multiple theories, data and methods has slowed progress towards making unified inferences of ecological change generalizable across large spatial, temporal and taxonomic scales. However, recent progress towards a theoretical synthesis now provides a guiding framework for organizing and integrating all primary data and methods for spatiotemporal assemblage-level inference in ecology. In this paper, we describe how recent theoretical developments can provide an organizing paradigm for linking advances in data collection and methodological frameworks across disparate ecological sub-disciplines and across large spatial and temporal scales. First, we summarize the set of fundamental processes that determine change in multispecies assemblages across spatial and temporal scales by reviewing recent theoretical syntheses of community ecology. Second, we review recent advances in data and methods across the main sub-disciplines concerned with ecological inference across large spatial, temporal and taxonomic scales, and organize them based on the primary fundamental processes they include, rather than the spatiotemporal scale of their inferences. Finally, we highlight how iteratively focusing on only one fundamental process at a time, but combining all relevant spatiotemporal data and methods, may reduce the conceptual challenges to integration among ecological sub-disciplines. Moreover, we discuss a number of avenues for decreasing the practical barriers to integration among data and methods. We aim to reconcile the recent convergence of decades of thinking in community ecology and macroecology theory with the rapid progress in spatiotemporal approaches for assemblage-level inference, at a time where a robust understanding of spatiotemporal change in ecological assemblages is more crucial than ever to conserve biodiversity.

Keywords: community ecology, dynamics, macroecology, paleoecology, processes, scale

NSO
NORDIC SOCIETY OIKOS

www.ecography.org

## Introduction

Understanding whether the processes underlying change in ecological assemblages – sets of species co-occurring in space and time – are generalizable across large spatial, temporal and taxonomic scales is a primary goal of macroecology (McGill 2019). Although macroecologists have identified a number of assemblage-level patterns frequently occurring in nature (Gaston and Blackburn 2000), attempts to infer general processes from those patterns using a 'pattern-first' approach have had only limited success because each pattern can often result from a variety of different causal pathways (Lawton 1999, Mayfield et al. 2010, Vellend 2016). Even when reliable pattern–process connections have been made, seldom have they led to general ecological theories, due to the high contingency of those connections on the spatial, temporal and taxonomic scale of study (Simberloff 2004, Vellend 2010).

Recent progress towards general theory has relied on a 'process-first' approach, which describes the fundamental processes underlying the patterns we observe in nature from first principles (Hubbell 2001, Alonso et al. 2006, Harte 2011, Marquet et al. 2014, Latombe et al. 2015, Vellend 2016). Ecological theories derived from first principles provide us with useful predictions of what the world should look like if only those fundamental processes were at play (Rosindell et al. 2012, Marquet et al. 2014). However, to become general, such theoretical predictions have to be repeatedly confronted with empirical patterns across spatial, temporal and taxonomic scales, and progress on this front has been slow (May 2004, Luo et al. 2011, Levin 2012, Mccarthy et al. 2012).

Ultimately, the main obstacle for both pattern-first and process-first approaches to macroecology has been the inability to examine patterns and processes across the entirety of spatial and temporal scales relevant to understanding assemblage changes (Levin 2012, Leibold and Chase 2018). Today, the availability of ecological data from disparate ecological sub-disciplines across spatial and temporal scales is increasing and a panoply of novel analytical methods have been developed to make use of these data; yet, the inability to fully integrate data and methods from different sub-disciplines remains a major obstacle to understanding the spatiotemporal dynamics of assemblages (Price and Schmitz 2016).

Discussion of integration among ecological data sources and methods has generally focused on how integration may elucidate the general processes underlying assemblage dynamics (Cavender-Bares et al. 2009, Beck et al. 2012, Fritz et al. 2013, Price and Schmitz 2016, Pearse et al. 2018). While this focus has led to the consensus that integration is a key step, it has been less successful at providing conceptual and practical ways in which integration among different spatiotemporal data sources and methods can be implemented. A powerful alternative approach could be to focus on how general processes cutting across spatial and temporal scales can provide a path to integration. There is already evidence that focusing on particular processes can facilitate integration among diverse spatiotemporal data sources and methods. For instance, a focus on body size-selective extinction in mammalian assemblages has resulted in the integration of disparate data sources across tens of thousands of years and across the globe to improve our understanding of change in past, contemporary and likely future extinction risk in mammals (Smith et al. 2018). Similarly, a focus on biotic interactions enabled the comparison of assemblage structure across 80 plant and animal assemblages spanning the past 300 million years (Lyons et al. 2016).

In this paper, we aim to foster integration among ecological sub-disciplines by highlighting how cross-scale processes relate to and connect disparate spatiotemporal ecological data sources and methods. First, we summarize the set of fundamental spatiotemporal processes driving change in assemblages by reviewing recent theoretical syntheses of community ecology, and develop a conceptual diagram that places those processes within a two-dimensional space–time 'process-space'. Second, we review recent advances in data and methods across the primary ecological sub-disciplines concerned with multispecies inference across both space and time, and associate each advance with the particular spatiotemporal process to which it is most relevant. Finally, we highlight how organizing data sources and methodological frameworks based on the primary fundamental processes they capture provides fruitful avenues for cross-pollination and integration among different data, methods and ecological sub-disciplines across large scales, thus furthering the mission of macroecology.

## Assemblage-level processes in space and time

What are the processes that cause the structure of ecological assemblages to change over space and time? While a single unified answer to this question is still missing, significant strides have been made towards a theoretical framework for understanding the dynamics of ecological assemblages. A key breakthrough was Vellend's conceptual synthesis of community ecology (Vellend 2010, 2016), building on prior work exploring the similarities between population genetic theory and community ecology (reviewed by Hu et al. 2006). Vellend's main intuition was to recognize that the myriad processes put forward by most major theories in community ecology can be grouped into just four fundamental 'high-level' processes (Vellend 2010, 2016): the movement of organisms across space (dispersal), random changes in species' relative abundances (drift), deterministic differences in fitness between individuals of different species (selection), and the creation of new species (speciation). According to this view, while proximate (or 'low-level') processes structuring communities might differ greatly among communities, differences in community structure are ultimately due to the influence of these four fundamental processes.

Vellend's synthesis does not claim to explain every single observation of a community or assemblage, and ecologists are likely to disagree on the relevance and importance of each fundamental process. However, this synthesis provides an invaluable theoretical basis for integrative and large-scale thinking in ecology. For instance, it has been widely adopted by microbial ecologists to understand the assembly of microbial communities (Hanson et al. 2012, Nemergut et al. 2013, Stegen et al. 2015, Evans et al. 2017, Zhou and Ning 2017); growing evidence that this framework can apply to microbes as well as macro-organisms highlights its applicability across large taxonomic scales. Moreover, the explicit consideration of speciation as a fundamental process provides a stronger link between ecological and evolutionary processes, as well as regional and local controls on assemblage dynamics. Finally, because ecological communities are defined independently from space or time (Vellend 2010, 2016), all four fundamental processes can apply irrespective of the spatial and temporal scale of study.

Here, we employ Vellend's conceptual synthesis – as well as complementary theoretical frameworks (Leibold and Chase 2018) – to delimit and visualize a 'process-space'

summarizing the high-level processes underlying change in assemblages across both spatial and temporal scales (Fig. 1). We start by considering a focal assemblage – defined as a group of co-occurring and potentially interacting organisms from multiple species – found at a given locality in the present (Fig. 1A). We then place the focal assemblage in its broader spatiotemporal context (Leibold et al. 2004, Leibold and Chase 2018) and depict how different processes may have acted 1) away in space from the focal locality (on the x-axis) and/or 2) backwards in time from the present (on the y-axis) to shape the structure of the focal community. Our conceptualization is reminiscent of that of (see Fig. 2 in Holt and Gaines 1992 ) who, in order to understand the trait composition of a given local population, invoked the concept of a 'phylogenetic envelope': 'the slice of space–time causally relevant to determining the abundance and composition of a population' (Holt and Gaines 1992). We frame the discussion here in terms of explaining present assemblage structure, but note that this framework also applies to focal assemblages that were observed in the past.

The focal assemblage is nested within the regional species pool or metacommunity (Fig. 1B), which is itself nested
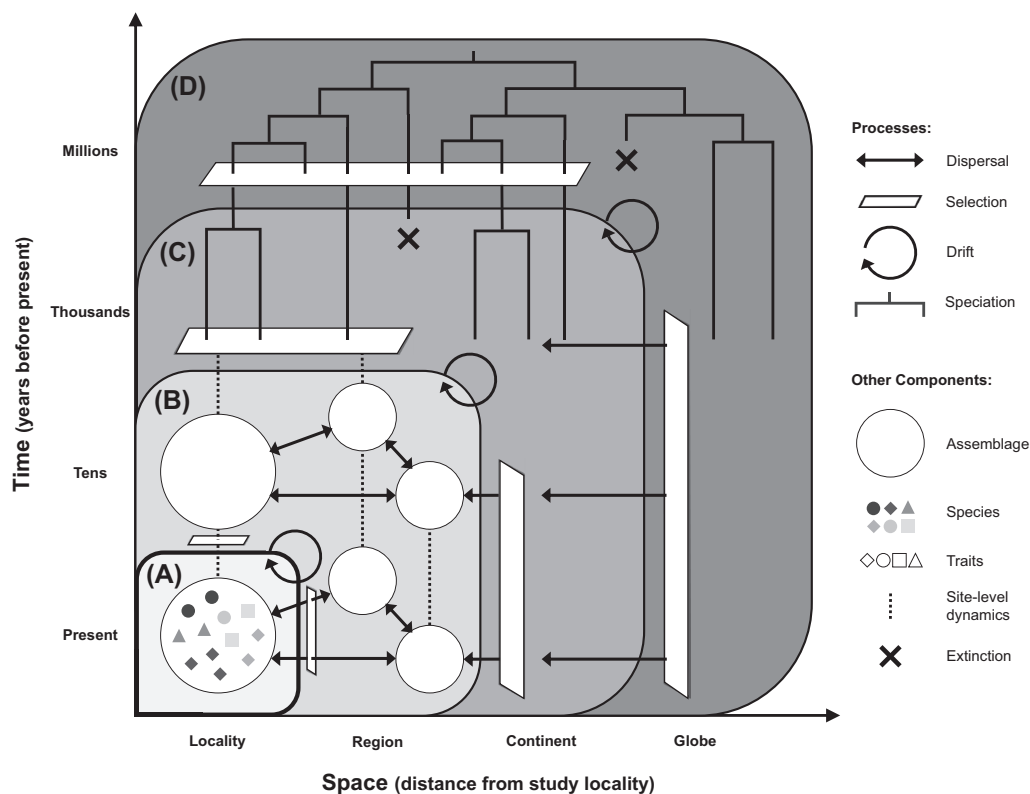
Figure 1. A process-space for understanding the dynamics of ecological communities across space and time. This conceptual diagram summarizes how the four fundamental (or 'high-level') processes included within recent theoretical syntheses in community ecology (Vellend 2010, 2016, Leibold and Chase 2018) – dispersal, drift, selection and speciation – influence the structure and dynamics of communities across across space and time. (A) The focal community: a group of co-occurring and potentially interacting organisms from multiple species and with different traits and degrees of phylogenetic relatedness occurring at present at a given locality. (B) The regional species pool over tens to hundreds of year before present from which the focal community arises. (C) The continental species pool over thousands to millions of years from which the regional species pool arises. (D) The global species pool over millions of years is the top level of the hierarchy.

within a larger continental species pool (Fig. 1C) and, in turn, within the global species pool (Fig. 1D). The processes of drift, dispersal, selection and speciation act across the entire spectrum of spatial and temporal scales to determine which individuals and species occur in the focal assemblage; the relative importance of these assemblage-level processes, however, is likely to differ based on the spatiotemporal position within this hierarchy. For instance, speciation is thought to occur mostly over longer timescales (Mittelbach and Schemske 2015; Fig. 1C–D), at least in communities of multicellular organisms (but see Lamichhaney et al. 2017). Indeed, when considering time frames long enough to allow for adaptation and speciation, the global and regional 'pools' of species that can colonize a focal locality become dynamic and local assemblage assembly increasingly is influenced by allopatric speciation and its underlying mechanisms (McPeek and Brown 2000, Mittelbach and Schemske 2015, Pigot and Etienne 2015, Patzkowsky 2017). Although Fig. 1 also depicts extinction, we do not consider extinction as a separate high-level process but rather as a natural outcome of selection and drift (following Vellend's framework; Vellend 2010, 2016).

In the following two sections, we review recent advances in ecological data and methods, with specific reference to how these enable us to infer and model the fundamental spatiotemporal assemblage-level processes of dispersal, drift, selection and speciation (Fig. 1).

## Inferring assemblage-level processes in space and time: data sources

### The starting point: multispecies occurrence data in space and time

One common starting point to infer empirically any of the four fundamental processes driving change in ecological assemblages across both space and time is a dataset comprising multispecies occurrence data across multiple locations and over multiple time periods. This starting point also serves to focus the scope of our review: many productive studies have focused on examining changes in assemblages at a very local scale using both observational data and experiments (Tilman et al. 1996, Crawley et al. 2005, Silvertown et al. 2006, Suttle et al. 2007), or alternatively examined drivers of changes in assemblages across space (Gaston 2000, Currie et al. 2004, Brown 2014, Fine 2015) or through time (Jackson et al. 2005, Blois et al. 2010, Rowe and Terry 2014, Terry and Rowe 2015). Here, we focus on the integration of space and time, which means that our approach and the key questions we are aiming to address are more macroecological in nature than within the realm of traditional community ecology.

Spatiotemporal assemblage data may be generated by several approaches. While replication over progressively larger spatial scales typically includes progressively larger aggregates of the same data types, replication over progressively larger temporal scales involves the use of fundamentally different data types. Here, we distinguish between three types of temporal data: contemporary, historical and paleo. While overlapping to some degree, these data types capture distinct temporal scales and traditionally are collected using different tools.

Contemporary data represent samples of multiple species generated by repeatedly surveying multiple sites at multiple times, often using a standardized sampling protocol. Datasets are typically aggregated from field surveys, censuses and monitoring schemes, which involve the direct or indirect detection/non-detection of live organisms. Consequently, contemporary data usually are the data source with the most precise spatial, temporal and taxonomic references. Since the mid-20th century, a growing number of long-term contemporary ecological research and monitoring networks have been providing detailed data on spatiotemporal assemblage change (Magurran et al. 2010, Lindenmayer et al. 2012, Haase et al. 2018). Some of these networks were set up directly following a priori funding and design [e.g. US Long-term Ecological Research Program (Hobbie et al. 2003); the National Ecological Observatory Network (Keller et al. 2008); Forestry Inventory and Analysis (Smith 2002); Breeding bird and butterfly surveys in North America (Swengel 1990, Sauer et al. 2017) and the UK (Roy and Sparks 2000, BTO et al. 2016, Brereton et al. 2017)], while others were assembled a posteriori, by facilitating the sharing and integration of information among studies and investigators gathering similar information [e.g. CTFS-ForestGEO (Anderson-Teixeira et al. 2015); GEOBON (Scholes et al. 2008)].

Historical data represent assemblage-level patterns over longer time scales, by coupling contemporary data with evidence of the occurrence of live species (e.g. museum specimens, written notes) collected in the same geographical area at some point in the past, often with a different purpose (Shaffer et al. 1998, Tingley and Beissinger 2009, Vellend et al. 2013). Typically, historical data are combined across multiple sites to reveal change over the spatial extent of a locality, state, country or region and the temporal extent of decades to a few centuries.

Historical data require a substantial amount of filtering and/or statistical treatment to account for sampling biases (Vellend et al. 2013), but recent advances are improving the potential of these data for understanding ecological community dynamics by minimizing the uncertainty in each datum's spatial, temporal and taxonomic reference. Much progress has occurred with respect to streamlining and improving approaches to georeferencing and digitization of museum specimens (Ellwood et al. 2015, Bloom et al. 2018) and physical maps derived from historical surveys (Kelly et al. 2016, 2017), integrating taxonomies and tracking taxonomic changes (Pyle and Michel 2008, Pyle 2016, Ytow 2016), and understanding the sources of uncertainty and bias in historical data sources across space (Rocchini et al. 2011, Meyer et al. 2015, 2016, Ruete 2015, Stropp et al. 2016),

time (Meyer et al. 2015, Stropp et al. 2016, Tessarolo et al. 2017) and taxa (Troudet et al. 2017). However, due to their largely opportunistic nature, historical data remain unreliable sources of information on species' relative abundances for many regions and times.

Paleo data enable the longest-term inferences of change. Our use of the term paleo refers to a variety of data types, including fossils and trace fossils, unified by their origin in soil or sedimentary deposits (Jackson and Blois 2015). Paleoecological data can be represented as either occurrence or abundance of different taxa, depending on the mode of collection (Maguire et al. 2015). The taphonomic processes underlying the deposition and representation of fossils within a sedimentary deposit dictate the spatial and temporal scales 'sensed' by a paleoecological dataset (Jackson 2012) and the certainty with which assemblage-level processes can be inferred. The bulk of paleo data are best suited to inferring processes operating at the landscape to regional scales over decades to millennia.

Recovery of paleoecological data has progressed along two main fronts. First, radiometric dating techniques and age-depth modeling have substantially improved our ability to situate these data in time. For example, radiocarbon dates can now be inferred with more accuracy and precision using smaller amounts of material (Bronk Ramsey et al. 2004, Bronk Ramsey 2008, Wood 2015) and calibration curves are continually being updated (Reimer et al. 2009, 2013). Moreover, new approaches to age-depth modeling provide better estimates of temporal certainty for unsampled portions of sedimentary deposits (Buck et al. 1999, Haslett and Parnell 2008, Bronk Ramsey 2009, Blaauw and Christeny 2011). Second, ancient DNA has been crucial for improving our ability to obtain species-level identifications from fossil assemblages (Heintzman et al. 2017) and extending the variety of taxa that could be sampled from sedimentary deposits (Pedersen et al. 2015).

One key to broad integration of data both within and across all three of the aforementioned data types is the development of common spatial, temporal and taxonomic metrics, standards and protocols (e.g. DarwinCore, <https://dwc.tdwg.org/>) that can allow these different data sources to 'talk' to one another (Brewer et al. 2012, Giovanni et al. 2012, Michener and Jones 2012, Parr et al. 2012, Hardisty and Roberts 2013, Kelly et al. 2017). Individual records – data from surveys of individuals, populations, communities – are often now integrated into databases, culminating in global databases of geo-referenced and time-referenced historical and contemporary specimens [e.g. iDigBio, <www.idigbio.org>; GBIF, <www.gbif.org> (Edwards 2004)] and fossils (Neotoma, <www.neotomadb.org> (Williams et al. 2018); PaleoBioDB, <paleobiodb.org>), as well as regional databases of historical vegetation maps [e.g. VTM, <vtm.berkeley.edu> (Kelly et al. 2016)] and photographs (e.g. CalPhotos, <calphotos.berkeley.edu>). Efforts to refine common data standards and geoinformatic infrastructure with a focus on linking data points from disparate sources based on shared spatial, temporal, and/or taxonomic attributes have

facilitated comparison of multispecies patterns among localities and times (Guralnick et al. 2007, Kissling et al. 2017).

Beyond the starting point of multispecies spatiotemporal occurrence data, a number of additional data types can help to infer the influence of different assemblage-level processes across large spatiotemporal scales. We discuss these ancillary data types and sources in the sub-sections below.

## Ancillary data on relative abundance: the promise of eDNA

The availability of relative abundance data, in addition to occurrence data, can improve our ability to infer key processes in ecology, particularly the importance of ecological drift (Magurran 2007, McGill et al. 2007, McGill 2010). While some of the data sources discussed above – primarily contemporary survey data and some types of paleo data – enable estimating the relative abundances of co-occurring species, our ability to do so across many taxa and repeatedly over space and time remains limited. Environmental DNA (eDNA; DNA captured from an environmental sample without first isolating any target organisms, Taberlet et al. 2012) provides a potential avenue for estimates of species' relative abundances across broad spatial, temporal and taxonomic scales (Bálint et al. 2018). Traces of eDNA may be recovered, for example, from faeces, skin cells and gametes, and may be sampled from contemporary environments (e.g. water, soil or air) or sedimentary deposits (Deiner et al. 2017). Although the exact relationship between abundance and the number of recovered sequences remains to be determined for many taxa and ecosystems, evidence so far indicates that eDNA samples include information on species' abundance ranks (Deiner et al. 2017). As a result, eDNA approaches for delineating community composition in contemporary and ancient samples are likely to soon provide complementary information to other sources of abundance and occurrence data (Bálint et al. 2018).

## Ancillary data for inferring dispersal

### Dispersal distances
Although indirect estimates of the process of dispersal can be derived from occurrence and/or abundance data alone (e.g. dispersal can be inferred from range size and shape measures (Whitmee and Orme 2012)), reliable estimates of the influence of dispersal on community-level patterns are vastly improved by direct measurements of dispersal distances (e.g. following birth or breeding) across multiple individuals of a focal taxon (Levine and Murrell 2003, Beckman and Rogers 2013). However, direct measurements of dispersal distances for different species are often lacking (Nathan et al. 2012). One promising avenue for obtaining measurements of individual-level dispersal is real-time animal tracking technology (Jønsson et al. 2016), though this approach is limited to a small subset of extant taxa. For most taxa, the best way forward may be to use existing experimental and empirical data on morphological, physiological, behavioral, and/or genetic

attributes of individuals, together with landscape-level attributes, to estimate the likely distribution of dispersal distances (Lischke et al. 2006, Hickler et al. 2009, Uriarte et al. 2009) and assess the sensitivity of community-level patterns to changes in dispersal parameters (Nathan et al. 2011, 2012, Beckman and Rogers 2013, Robledo-Arnuncio et al. 2014).

A major challenge in assessing the importance of dispersal is the accurate reflection of long-distance dispersal events (Trakhtenbrot et al. 2005, Nathan 2006, Nathan et al. 2012). The paleoecological record has and continues to be instrumental in furthering our understanding of potential long-distance dispersal events and their consequences on the dynamics of ecological communities (Clark et al. 1998, 2003, McLachlan and Clark 2004), although there are often discrepancies between dispersal rates estimated from paleontological and contemporary records (Clark 1998). Moreover, range shifts observed through the fossil record reflect, in part, the outcome of dispersal events and can provide indirect evidence of dispersal events (Ordonez and Williams 2013).

### Ancillary data for inferring selection

#### Abiotic environment
Linking community dynamics to spatiotemporal variation in the surrounding abiotic environment is the primary approach for inferring selection from abiotic processes. The rapid advancement of environmental sensor technologies and the advent of wireless sensor networks now enable taking physical and chemical measurements at temporal and spatial scales previously difficult to cover (Porter et al. 2012). These on-the-ground measurements are being integrated with additional historical and contemporary environmental data such as digital elevation models to generate openly-available high-resolution gridded products of variation in temperature, precipitation and moisture over the last 50–100 years (Kriticos et al. 2012, Kearney et al. 2014, Fick and Hijmans 2017, Karger et al. 2017, Vega et al. 2017, Abatzoglou et al. 2018). Moreover, regional to global simulations of different aspects of climate can now also be generated as gridded data for any period during the last 21 000 years and at temporal resolutions as high as a decade (Fordham et al. 2017). Finally, remote sensing data provide large amounts of near-real-time information on hydrology, geology, land cover/use and habitat availability, and are increasingly easy to access using platforms such as Google's Earth Engine (<earthengine.google.com>; Gorelick et al. 2017) or USGS BISON (<bison.usgs.gov>).

#### Biotic relationships
Although co-occurrence data can be used to infer whether a given pair of species potentially interact (Blois et al. 2014), inferences about the influence of selection through biotic processes are improved by direct estimates of the type and magnitude of interaction among co-occurring species and trophic links. These data may come from a variety of sources, including direct observations (Ballantyne et al. 2017), gut contents (Pompanon et al. 2012), and isotope analyses (Montoya et al.

2015), and are being gathered from many studies in open online repositories (Poelen et al. 2014).

#### Traits
Whether through abiotic or biotic factors, selection shapes the structure and dynamics of ecological communities by acting on ecological similarities and differences among organisms (Cadotte et al. 2011, 2013). Therefore, our understanding of selection has been greatly enhanced by examining variation in a range of ecological traits among individuals and/or species. As a result, traits are increasingly documented within centralized data repositories for broad taxonomic groups [e.g. TRY, (Kattge et al. 2011); BiolFlor, <www.biolflor.de>; PanTHERIA, (Jones et al. 2009); BIEN, (Enquist et al. 2009)]. For the most part, these data repositories provide species-level trait summaries and only rarely include estimates of individual-level trait variation. The availability of individual-level trait data remains a large gap to be filled (Pearse et al. 2018).

#### Extinction
Extinction is one possible outcome of selection. Therefore, determining where, when and how often extinctions have taken place can provide information on the potential cause of a selection event. Information on recent species extinctions and the likelihood of future extinction can be garnered from the International Union for the Conservation of Nature (IUCN) Red List of Threatened Species (IUCN 2019). However, by far the richest empirical record of extinction is contained within the fossil record (Quental and Marshall 2010, Barnosky et al. 2011).

### Ancillary data for inferring speciation

#### Speciation
Speciation is the only fundamental process, together with dispersal, that can add species to an assemblage (Vellend 2016). As a result, understanding how frequently speciation events occur is fundamental to assess the importance of speciation on differences in assemblage structure (Fine 2015, Hubert et al. 2015). Molecular phylogenies, which represent hypotheses of speciation events among species in the form of bifurcating trees (Pagel 1999), can be used to estimate speciation rates [but only after applying a number of assumptions (e.g. an accurate macroevolutionary process model) which may not always be warranted, Stadler and Bokma 2013]. Reliably estimating speciation rates has also long been a focus of paleobiology, and the fossil record includes high-quality information about the incidence and rate of speciation through time (Foote 2000, Sepkoski 2007, Mayhew et al. 2008).

## Modeling assemblage-level processes in space and time: advances in methods

Expansion in the types of readily available spatiotemporal ecological data has resulted in a corresponding boom

in analytical methods for making use of those data. In this section, we review recent advances in multispecies spatio-temporal ecological methods and how they can model drift, dispersal, selection and speciation. In doing so, we map out the relationships among processes, data types and methods in Fig. 2. Our mapping focuses on the data types that are most typically used in conjunction with a given method and process, and does not imply that particular methods are unable to incorporate other data types or model other processes. Furthermore, we do not aim to provide an exhaustive list of all existing analytical methods for multispecies spatiotemporal ecological data. Instead, we highlight methodological advances that show particular promise for integrating datasets spanning large spatial and temporal scales.

## Modeling drift

### Neutral individual-based models

Because the dynamics of assemblages emerge from the behavior of individual organisms, understanding assemblage-level properties ultimately benefits from incorporating the interactions among individuals and their

attributes. Consequently, individual-based models (IBMs) that model regional species pools from the bottom up have advanced our understanding of assemblage-level processes across space and time (Chave 2004, Rosindell et al. 2012, DeAngelis and Grimm 2014, Weathers et al. 2016). IBMs are composed of discrete individual organisms interacting with each other and their surrounding environment (DeAngelis and Grimm 2014). Each individual is assigned a set of state variables (e.g. spatial locality), attributes (e.g. physiological traits) and/or behaviors (e.g. dispersal, reproduction), which may vary among individuals and through time (Grimm et al. 2006).

Individual-based models of neutral dynamics (or neutral IBMs) with equal per capita birth and death rates (i.e. drift) and equal likelihood of colonization of sites (i.e. dispersal) among individuals of different species (Hubbell 2001, Rosindell et al. 2011) represent a starting point for more complex models of assemblage dynamics (Chave 2004, Rosindell et al. 2012). Empirical evaluation of predictions from these and related classes of neutral models have provided tests of the importance of ecological drift in shaping assemblage dynamics over decadal scales (Chave 2004,
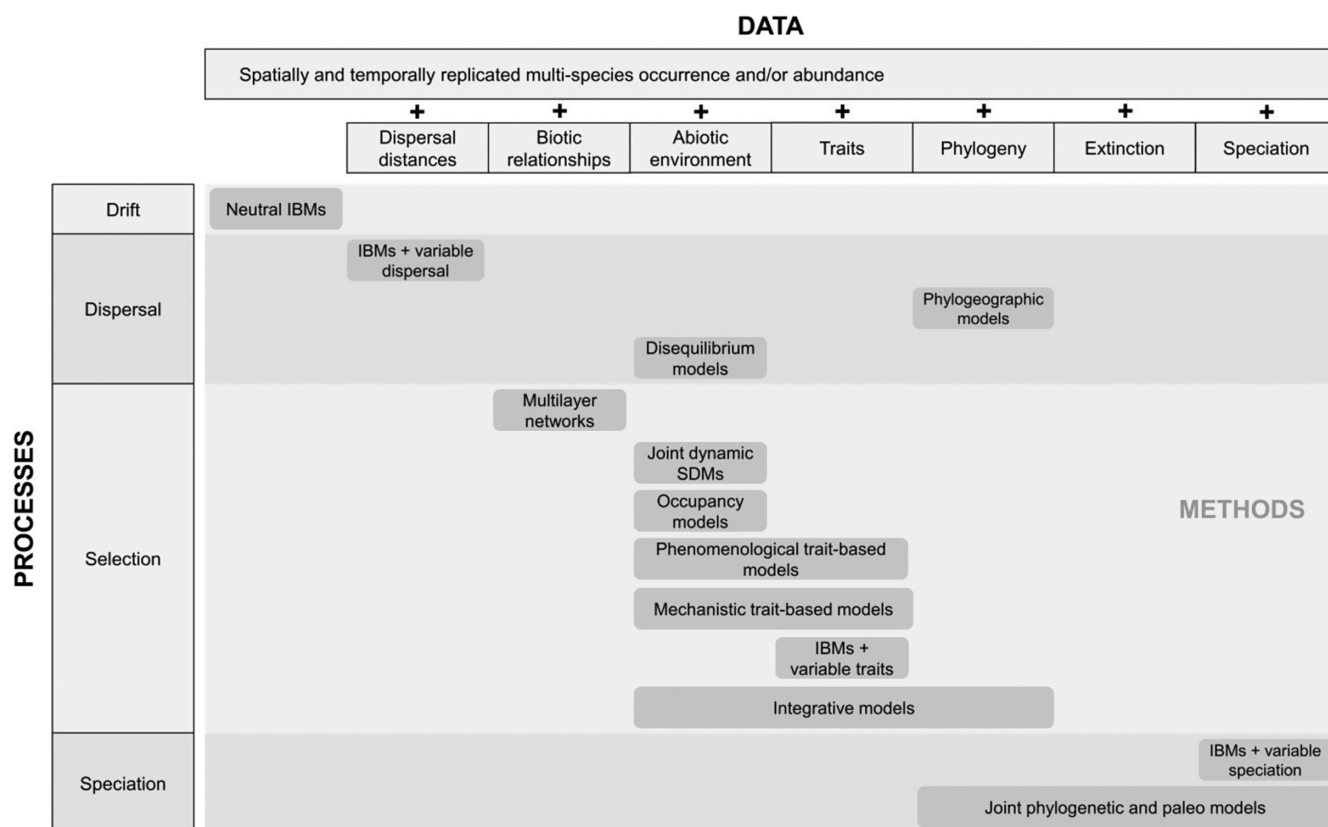
Figure 2. The relationship between processes, data types and methodological advances in the study of ecological community dynamics in space and time. Rounded grey boxes represent methodological advances discussed in the 'Modeling community processes in space and time: advances in methods' section. Each method is linked with the primary community-level process it models, as well as the main data types it uses to model the process. All methods make use of at least spatially and temporally replicated multi-species occurrence and/or abundance data, and most methods also make use of one or more additional data types. Abbreviations: IBMs = individual-based models; SDMs = species distribution models.

Fuller et al. 2005, McGill et al. 2006a, Horvát et al. 2010, Rosindell et al. 2011). In rare cases, predictions from neutral IBMs have also been evaluated against paleoecological data at millennial scales (McGill et al. 2005, Sclafani and Holland 2013, Correa-Metrio et al. 2014, Holland 2018). We are unaware of studies evaluating predictions from neutral IBMs against integrated datasets of paleo and historic or contemporary communities; however, such a study could enable assessing the importance of ecological drift at both large temporal extents and fine temporal resolutions.

## Modeling dispersal

### Individual-based models with variable dispersal among species
IBMs often assume identical dispersal abilities among species due to the lack of data to parameterize differences in dispersal kernels (Nathan 2006, Nathan et al. 2011, Robledo-Arnuncio et al. 2014). However, IBMs that break neutrality in dispersal through the introduction of variable dispersal ability among species can provide a better fit to empirical data than neutral models (Jabot 2010, Janzen et al. 2015). Although IBMs with variable dispersal have rarely been applied to the fossil record, evaluating predictions from IBMs whose dispersal distributions have shorter or longer right tails against paleoecological data over millennia could help refine estimates of the incidence of long-distance dispersal events (Clark 1998).

### Phylogeographic models
Phylogeographic models represent another approach to identify the assemblage-level outcome of long-distance dispersal events. These models rely on combining molecular phylogenies with ancestral state reconstructions to identify the origin of a focal set of taxa and their subsequent dispersal to other regions (Marske et al. 2013). Testing how well phylogeographic models including or excluding species' long-distance events explain the current emergent pattern of species cooccurrence within a focal clade can enable understanding the assemblage-level effects of dispersal (Pennington and Dick 2004, Dick et al. 2007, Lam et al. 2018). More recently, the advent of approximate Bayesian computation (ABC) methods and high-throughput DNA sequencing methods have revolutionized our ability to directly test the effects of different dispersal processes on the aggregate distribution of cooccurring species (Hickerson et al. 2010, Chan et al. 2014). One particularly exciting avenue is the combined use of phylogeographic models with the fossil record to assess the concordance of multiple lines of evidence of past dispersal events, though this avenue remains underexplored at the assemblage level (Hickerson et al. 2010, Metcalf et al. 2014).

### Disequilibrium models
The influence of dispersal can also be inferred indirectly using models that correlate multispecies spatiotemporal occurrence data with the surrounding abiotic environment, with the objective to identify the imprint of dispersal limitation on assemblage dynamics. For instance, contemporary distribution patterns are often better explained by late Quaternary climate than by contemporary climate, indicating that assemblages are not in equilibrium with contemporary climate, possibly due to dispersal limitation (Svenning and Skov 2007, Svenning et al. 2008, Kissling et al. 2016). In addition, recently developed approaches can measure the extent of mismatch between assemblages and their environmental context (Blonder et al. 2015) and classify assemblage dynamics under climate change into different equilibrium categories (Blonder et al. 2017). These approaches rely on assumptions of equilibrium between an assemblage and climate as a way to assess the extent of assemblage disequilibrium with climate. With the increasing temporal resolution of paleoecological and paleoclimate data, disequilibrium models could be used to assess fluctuations in the influence of dispersal limitation over millennia.

## Modeling selection

### Trait-based models
Recent advances in trait-based models are uncovering direct links between trait variation, environmental gradients and organismal fitness, the key to understanding the influence of selection on assemblage dynamics (Cornwell et al. 2006, Shipley et al. 2006, Savage et al. 2007, Suding et al. 2008, Dray and Legendre 2008, Messier et al. 2010, Shipley 2010, Mason et al. 2013, Brown et al. 2014, Warton et al. 2015a, Loranger et al. 2016, Ovaskainen et al. 2017). Trait-based models may involve empirically relating intra- and/ or inter-specific trait values to environmental variation using correlative models and projecting how the spatio-temporal distribution of traits and their performance across environmental gradients give rise to assemblage-level properties (Laughlin et al. 2012, Lasky et al. 2014, Laughlin and Messier 2015). Alternatively, trait-based models may involve describing trait–environment relationships mechanistically from first principles (Norberg et al. 2001, Shipley et al. 2006, Savage et al. 2007, Jabot 2010, Enquist et al. 2015, Holt and Chesson 2016, Scherer et al. 2016, Galic et al. 2018). Such mechanistic models are highly theoretical and translate common notions of trait-based mechanisms into dynamic differential equations.

Trait-based modeling advances can improve our ability to model selection by incorporating ecological differences among individuals (Cadotte et al. 2013). However, whether these models generate reliable and general inferences of complex community dynamics across large spatial and temporal scales remains to be seen. Exploiting the wealth of information on morphological and ecological trait variation held in the fossil record (Blois et al. 2013), in conjunction with downscaled paleoclimate data, may provide much-needed validation for trait-based model predictions. In cases where only a few traits are available or trait data are missing altogether, the use of phylogenetic distances among taxa may represent a valid alternative, under the assumption that closely related species have more similar traits than do distantly related species (Cavender-Bares et al. 2009). More recently,

the combined use of trait and phylogenetic distances has further improved our ability to infer the processes shaping assemblages (Cadotte et al. 2013, Cadotte and Davies 2016, Gianuca et al. 2017).

### Joint dynamic species distribution models
A fundamental challenge when exploring the assemblage-level consequences of selection is to disentangle the influence of selection through biotic versus abiotic factors. This is because both abiotic and biotic processes may result in largely similar assemblage-level patterns (Gross et al. 2009, Mayfield and Levine 2010). Moreover, there may be causal interdependencies and feedbacks between abiotic and biotic factors (Dunson and Travis 1991, Suttle et al. 2007), such that biotic factors (e.g. competition) influence physiological tolerances to abiotic factors (e.g. cold-tolerance), and vice versa. One way to begin tackling these challenges at large spatiotemporal scales is to use joint species distribution models (SDMs). These models make use of pairwise species co-occurrence matrices to jointly estimate how the distributions of multiple species are associated with each other, as well as with their abiotic environment (Dorazio 2014, Pollock et al. 2014, Fithian et al. 2015, Warton et al. 2015b, Taylor-Rodriguez et al. 2017). Joint SDM frameworks and related null model approaches based on co-occurrence data have successfully been applied to both contemporary (Pollock et al. 2014, Fithian et al. 2015) and paleoecological data (Blois et al. 2014, Lyons et al. 2016, Smith et al. 2016). However, simple formulations of these models should be used with care as they have been found to capture some biotic interaction types unreliably (Zurell et al. 2018).

Recently, joint SDMs have been extended to incorporate temporal (i.e. joint dynamic SDMs), as well as spatial, processes by estimating how the relative abundance of each species may enhance or inhibit the subsequent abundance of itself and/or other species in the community (Thorson et al. 2016, Chesson 2017, Ovaskainen et al. 2017, Schliep et al. 2018). In addition, because interspecific interactions are likely to vary in direction and strength under different environmental conditions, joint species distribution models have also been extended to incorporate environmentally dependent biotic associations (Tikhonov et al. 2017). A comprehensive evaluation and comparison of many of these SDM approaches has recently been published (Norberg et al. in press); we refer the reader to it for more detailed information on these models.

### Multilayer networks
Biotic interactions are more complex than positive or negative associations of constant magnitude among pairs of co-occurring species: they differ in type, direction and magnitude across multiple trophic levels and in space and time. To tackle this complexity, progress is being made towards the integration of ecological and trophic networks within dynamic and spatially-explicit meta-community models (Poisot et al. 2015, Gravel et al. 2016, Pilosof et al. 2017, Coelho and

Rangel 2018, Godoy et al. 2018). One way to do this is to use ecological multilayer networks (Pilosof et al. 2017): networks of nodes connected not only to other nodes within the same network but also to the same nodes across different 'layers' (e.g. different locations and/or points in time). Additionally, incorporating mutualistic networks within spatially-explicit individual-based models of stochastic meta-communities can reliably reproduce the structural patterns observed in realized pollination and seed dispersal networks (Coelho and Rangel 2018). Although fewer studies have quantified the importance of biotic interactions over centennial to millennial scales, ecological network reconstructions inferred from variation in functional morphology or stable isotope ratios within the paleoecological record represent a potential avenue for examining changes in biotic interactions at large spatial and temporal scales (Domingo et al. 2012, Blois et al. 2013, Yeakel et al. 2014).

### Hierarchical occupancy models
Regardless of the data collection approach, all assemblage-level data sources are subject to the imperfect detection of individuals at the focal locality and time. While the causes of imperfect detection may differ greatly between contemporary, historical and paleo community approaches, the consequences are similar: failure to detect individuals of common species results in underestimation of abundance; failure to detect individuals of rare species results in false absences (Iknayan et al. 2014, Guillera-Arroita 2017). Hierarchical occupancy models have become the standard way to account for imperfect detection in analyses of contemporary ecological data (Royle et al. 2007, Dorazio et al. 2010, Bailey et al. 2014, Iknayan et al. 2014, Guillera-Arroita 2017). They do so by separating and independently estimating the ecological process of interest from the observation or data-generation process, which is contingent on the sampling protocol and subject to imperfect detection (Bailey et al. 2014). Hierarchical occupancy models are now able to model changes concurrently over space and time (Dorazio et al. 2010, Yamaura et al. 2011, Van Strien et al. 2011, Van Strien et al. 2013, Bailey et al. 2014, Isaac et al. 2014) and accommodate assemblage-level data to estimate the probability of missing species (Iknayan et al. 2014, Pacifici et al. 2017).

Owing to their ability to account for different types of sampling bias, occupancy models are also becoming key to modeling longer-term data sources. For instance, the application of hierarchical occupancy models to museum specimens and historical observations can vastly improve our estimates of the influence of selection on assemblages at the centennial scale (Tingley and Beissinger 2009, 2013, Tingley et al. 2012, Iknayan et al. 2014, Rowe et al. 2014, Rapacciuolo et al. 2017, Zeilinger et al. 2017). Although different approaches have traditionally been preferred to account for imperfect detection in paleoecological data (Alroy et al. 2001, Sugita 2007a, b, Pirzamanbein et al. 2014), hierarchical occupancy models are beginning to be applied to the fossil record for assemblage-level inference (Liow 2013).

With respect to data integration, one of the most appealing features of hierarchical occupancy models is that they can account for multiple data-generation processes within the same hierarchical model (Dorazio 2014, Fithian et al. 2015). This shows tremendous promise for accommodating different spatiotemporal data sources with different underlying biases in a single model, thus extending the generality of spatiotemporal inferences.

### Integrative models

The past several years have seen the development of integrative analytical frameworks, which combine several of the theoretical and methodological advances discussed above. Among those, two recent frameworks – hierarchical modelling of species communities (HMSC; Ovaskainen et al. 2017) and generalized joint attribute modeling (GJAM; Clark et al. 2017) – combine recent advances in joint species distribution models and the estimation of imperfect detection from multiple data sources with trait– and phylogeny–environment relationships, in an attempt to model the influence of multiple abiotic and biotic selection filters. GJAM is a modeling framework aimed at integrating different types of observational data and/or environmental attributes collected during ecological surveys. This approach has been used to explore the structure of vegetation assemblages by integrating species-level and plot-level data (Clark et al. 2017), determine the response of the soil microbiome to environment (Clark et al. 2017), and identify the importance of arbuscular mycorhizzae fungi across successional stages of tropical trees (Bachelot et al. 2018). HMSC is perhaps even more tailored for use with assemblage-level data of different spatiotemporal scales, in that it explicitly uses information on species niches, phylogenetic relationships and traits, and random variation and covariation among species occurrences across space and time, to estimate the combined influence of biotic and abiotic selection on assemblage-level patterns (Ovaskainen et al. 2017).

Another development in integrative modeling – scale transition theory – shows particular promise for integrating across spatial and temporal scales (Melbourne and Chesson 2005, Chesson 2009, Holt and Chesson 2016). Scale transition theory predicts the emergent behavior of metacommunities at large spatial and/or temporal scales from the smaller spatial- and/or temporal-scale behavior of populations. Scaling up relies on quantifying how nonlinearities in population dynamics at smaller scales vary with spatial and/or temporal variation in the environment at larger scales (Melbourne and Chesson 2005). These models are integrative because they estimate parameters summarizing nonlinearities at the local scale from experimental data and combine these with spatial and/or temporal variation estimated from field sampling programs, for example, to identify how properties change or emerge as the scale expands. Though we have included this approach in the 'Modeling selection' section, the method also includes aspects of other processes as well, particularly dispersal (Holt and Chesson 2016).

Owing to their flexibility in accommodating various processes and data types, integrative models such as HMSC, GJAM and scale transition theory may provide unified frameworks for the study of selection on multispecies dynamics across large spatiotemporal scales.

## Modeling speciation

### Individual-based models with variable speciation rates among species

One way to assess the influence of speciation on assemblage dynamics is to break neutrality in speciation within IBMs. For instance, incorporating different modes and rates of speciation within IBMs has revealed complex relationships between speciation, richness and the spatial organization of populations of different species (Davies et al. 2011, Desjardins-Proulx and Gravel 2011, 2012, Hubert et al. 2015).

### Joint phylogenetic and paleoecological models

Geographical variation in the relative abundance of different taxa and guilds within the fossil record has traditionally been used to reveal how species origination, as well as extinction, in one region may have influenced species diversity in another (Jablonski et al. 2006, 2017, Patzkowsky 2017). In parallel, molecular phylogenies have been used to examine how geographical variation in the rates of speciation has influenced assemblage-level patterns such as species, functional and/or phylogenetic diversity (Belmaker and Jetz 2015, Marin and Hedges 2016, Oliveira et al. 2016, Marin et al. 2018) and to assess patterns of phylogenetic structure within and among regional species pools (Lososová et al. 2015, Pigot and Etienne 2015, Lessard et al. 2016, Carlucci et al. 2017). More recently, fossils and molecular phylogenies are being combined within joint phylogenetic and paleo models. These models either place extinct taxa within a backbone molecular phylogeny or jointly estimate tree topology and divergence times to improve estimates of speciation and diversification rates and their subsequent influence on communities (Ezard et al. 2011, 2012, Morlon et al. 2011, Etienne et al. 2012, Condamine et al. 2013, Pennell and Harmon 2013, Lamsdell et al. 2017).

## Towards an integrative approach for understanding assemblage change in space and time

We are far from the first to discuss integration among approaches in community ecology (Beck et al. 2012, Fritz et al. 2013, Price and Schmitz 2016, Pearse et al. 2018). Our perspective on integration is guided by recent progress towards a unified theory of ecological communities (Vellend 2010, 2016, Leibold and Chase 2018; see also Pásztor et al. 2016, Worm and Tittensor 2018) and our view is that one of the major barriers to integration is conceptual: the seemingly intractable disparity and complexity – in processes, data

and methods – that underlies ecological systems at large spatial, temporal and taxonomic scales. Yet, although their specific objectives and tools may differ greatly, all approaches to assemblage-level inference in space and time may be thought of as ultimately guided by the same overarching question: what is the relative importance of the processes of drift, dispersal, selection and speciation on focal assemblages given their spatial and temporal context?

Despite the emergence of this process-based framework (Fig. 1; Vellend 2010, 2016), disentangling the relative importance of all four processes across all scales of space and time remains a challenge. To date, much of the progress has been made by attempting to elucidate most or all processes within the spatiotemporal bounds of a particular sub-discipline. We demonstrate how the complementary approach of focusing on only one (or a subset) of all potential community-level processes, but across all relevant scales of space and time, may provide an effective avenue to overcome the conceptual barriers to integration among disparate data and methods.

We argue that the degree of methodological complexity and data integration required for most multispecies spatiotemporal ecological studies could be established using the following sequential steps (Fig. 3): 1) focus on one of the four fundamental processes (drift, dispersal, selection, speciation) which is likely to influence the structure of the focal assemblage and whose influence has yet to be explored; 2) identify how many steps in time and space are likely to be required to sufficiently quantify the influence of this process on assemblage-level change; 3) identify the necessary biological units of analyses (individual, population, species, community); and 4) identify the primary and ancillary data sources and methods required to infer the influence of this

fundamental process in this system (Fig. 2). This cycle of four steps could then be repeated sequentially for each of the four fundamental processes (Fig. 3), resulting in a comprehensive exploration of the primary processes underlying spatiotemporal assemblage change in four sets of analyses. In Box 1, we provide a case study examining the influence of dispersal on small mammal assemblages in California's Sierra Nevada mountains to illustrate how these four steps can help organize an integrative macroecology study.

Of course, there are practical barriers to the implementation of this organizing paradigm. First, many researchers are making substantial contributions within their core sub-discipline, yet progress towards integrating data and methods across space and time requires focusing on the spatiotemporal boundaries between sub-disciplines. For example, re-surveys and other approaches in ecology have been successful at linking historical with contemporary communities (Tingley and Beissinger 2009, Rapacciuolo et al. 2014a, Goring et al. 2016) but fewer approaches exist to align the recent paleo record with historical and contemporary records (Seddon et al. 2014, Jackson and Blois 2015; but see Terry 2010, Tomašových and Kidwell 2017). Even within paleo approaches, schisms emerge between work done in more recent times (i.e. the latest Pleistocene and the Holocene) and older times in Earth history. While some research groups are advancing methods for integrating data across the spatiotemporal boundaries of sub-disciplines, they are not often focused specifically on a process-based approach to community ecology.

Second, in order to make connections across boundaries in practice, different sub-disciplines need to be able to talk to one another. We are making rapid progress towards integration of assemblage-level data across disparate spatial and temporal scales through the use of common formats with standardized spatial, temporal and taxonomic units and estimates of uncertainty (Guralnick et al. 2007, Kissling et al. 2017). Nevertheless, even when standardized, metrics based on species identity are problematic at very large spatial and temporal scales due to the dynamism of the global and regional species pools over long timescales (Mittelbach and Schemske 2015). Therefore, identifying common currencies that allow cross-talk across these large scales will be key (Barnosky et al. 2017). Traits – whether phenotypic, ecological or metabolic – have gained traction as the 'default' common currency (McGill et al. 2006b, Eronen et al. 2010, Violle et al. 2014, Enquist et al. 2015, Funk et al. 2017). However, other attributes of individuals, communities, localities or time periods may also provide equally useful common currencies for examining ecological communities across scales; these include area- or taxon-based climate velocity vectors (Loarie et al. 2009, Tingley et al. 2012), ecological network metrics (Heleno et al. 2014, Toju et al. 2017), species- or assemblage-level environmental equilibrium measures (Svenning and Skov 2004, Rapacciuolo et al. 2014b, Svenning et al. 2015, Blonder et al. 2017), or estimates of community novelty (Urban et al. 2012, Graham et al. 2017).
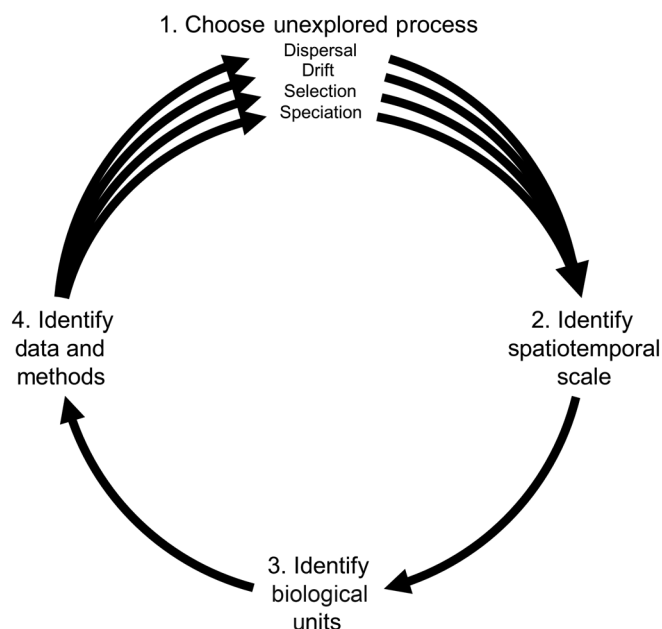
Figure 3. An organizing paradigm to frame spatiotemporal community ecology research studies and programs using four sequential steps.

**Box 1. Integrating multispecies spatiotemporal data and methods in light of theory: a case study**

*An organizing paradigm for macroecological studies spanning large spatial, temporal and taxonomic scales*
Recent theoretical developments in community ecology now provide an organizing paradigm for integrating advances in data and methods across disparate ecological sub-disciplines and across large spatial and temporal scales. This organizing paradigm can be summarized as a cycle including four sequential steps (Fig. 3). Each iteration of this cycle aims to elucidate the influence of one of four fundamental spatiotemporal assemblage-level processes, and multiple iterations of this cycle – at least one for each process – enable a comprehensive exploration of the suite of processes underlying spatiotemporal assemblage-level change in a focal community. Approaching studies in this way will help establish and subsequently overcome the necessary complexity and integrative nature of macroecological studies at large spatial, temporal and taxonomic scales. Below, we work through one iteration of the cycle using a case study of Sierra Nevada small mammal assemblages, focusing on the ideal attributes of such a study and recognizing that we may not reach that ideal in many (or most) cases.

*Step 1. Choose an unexplored process*
Given the focus on large scales of space and time, the first step is to narrow the focus of study to one of the fundamental spatiotemporal assemblage processes, thus making integration more manageable from a conceptual point of view. For instance, we might begin by quantifying the influence that the process of dispersal has had on the structure of contemporary small mammal assemblages in the Sierra Nevada mountains of California. The key question here is: have dispersal differences among species contributed to variation among assemblages across the Sierra Nevada?

*Step 2. Identify the most relevant spatiotemporal scale*
The second step is to identify the most appropriate spatiotemporal extent and resolution for quantifying the influence of the process chosen in step one on the focal assemblage In Fig. 1, we provide a broad overview of the spatiotemporal scales over which each spatiotemporal assemblage process is likely to be most relevant, though we emphasize that the choice of an appropriate spatiotemporal scope will depend on both the question being addressed and the study system. For example, for many questions addressing how dispersal dynamics among species have contributed to variation among assemblages at a relatively local scale, a temporal extent of decade(s) may be appropriate for some study systems but may need to be longer (centuries) for other study systems (e.g. long-lived trees). To examine variation among assemblages at the regional spatial scale and above, however, a temporal extent of millennia may be more appropriate even for short-lived taxa. For example, there is evidence of legacy impacts of Pleistocene climate on contemporary patterns of diversity in trees (Svenning and Skov 2007, Svenning et al. 2008, Kissling et al. 2016), potentially due to dispersal limitation, and Guralnick (2006) found an interaction between niche dynamics (i.e. selection) and recolonization history (i.e. dispersal) for North American small mammals, indicating that orbitally-forced range dynamics may be important to consider. Thus, the temporal extent necessary to understand the impact of dispersal on variation in assemblage structure is potentially on the order of 1000s of years for many study systems. In the case of Sierra Nevada small mammals, then, to understand the influence of dispersal on our example assemblages, the study might focus on a spatial extent including western North America and a temporal extent ranging from the Last Glacial Maximum (~21 000 years ago) to the present (Fig. I).

*Step 3. Identify the most relevant biological units*
The third step is to identify the most appropriate biological unit of analysis (i.e. individual, population, species, community). To quantify the influence of dispersal on assemblage structure across space and time, we need at the very least a species-level estimate of the mean dispersal event for each species in the assemblages; ideally, however, we would also have an estimate of the individual-level variation around this mean and, particularly, reliable estimates of the right tail of the dispersal probability density function (i.e. the probability and distance of long-distance dispersal events).

*Step 4. Identify the most relevant data and methods*
The fourth step is to identify the spatiotemporal assemblage data and methods most useful to answer the question specified in steps one to three. Fig. 2 summarizes the data and methods most likely to be useful for quantifying the influence of dispersal. To return to our case study, spatiotemporal occurrence data on small mammals in the Sierra Nevada over the last 21 000 years are available from a number of complementary sources (Fig. I). These include (but are not limited to): paleocommunity ecology data from Faunmap (FAUNMAP Working Group 1994), Neotoma (Williams et al. 2018), and individual long-term sedimentary records (Furlong 1904, 1906, Feranec et al. 2007, Blois et al. 2010); historical data from historical ecological surveys (e.g. Grinnell Resurvey Project; (Rowe et al. 2014)) and museum collections (e.g. Arctos, <https://arctosdb.org/>); and contemporary data from NEON sites occurring south and north of the focal region (i.e. the Soaproot Saddle and San Joaquin Experimental Range sites) and repeated ecological monitoring as part of individual studies (Kelt et al. 2013). In addition, ancillary data particularly useful for examining the influence of dispersal include estimates of the past rate and velocity of range shift (Williams and Blois 2018), modern estimates of traits related to dispersal ability (e.g. PanTHERIA, Jones et al. 2009, or Whitmee and Orme 2012), or dispersal events recorded directly. Assessment of the completeness of the data given the research questions may motivate additional data collection focused on filling gaps in assemblage data available for particular places and/or times, or designing experiments or new data campaigns aimed at reducing uncertainty in key parameters. The spatiotemporal methods most useful to examine dispersal include individual-based models, disequilibrium models and phylogeographic models (Fig. 2).
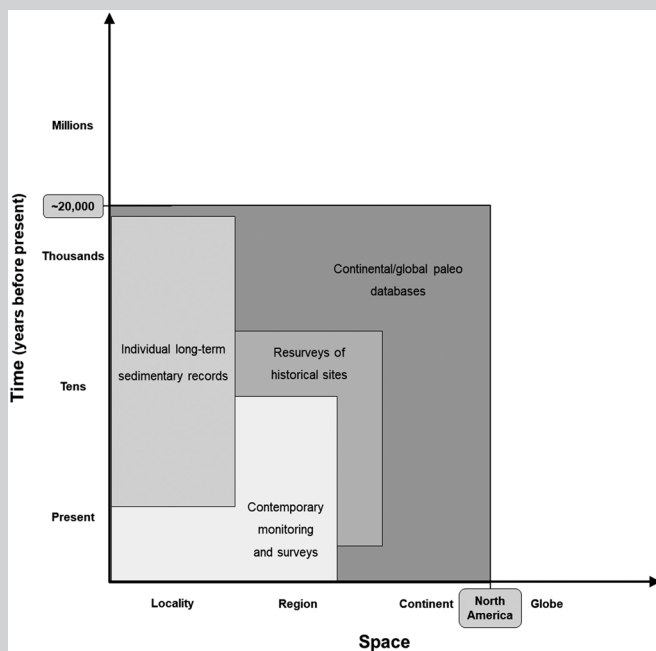
Figure I. Spatiotemporal scales and available data sources most relevant to a study of small mammal community structure across the California Sierra Nevada.

*Two potential iterations of the small mammal case study*
Given the multitude of data sources and methods relevant to any given study question, combined with the fact that we are often not able to attain 'ideal' given available data and methods, repeated iterations of steps 2–4 in Fig. 3, following the choice of a particular process, are expected to sequentially refine understanding of the focal process. For example, each iteration may result in a different analysis, with different iterations representing independent lines of evidence that help interpret the influence of the focal process on assemblage dynamics. Below, we present two potential iterations of the steps 1–4 outlined above with respect to our case study of the influence of dispersal on Sierra Nevada small mammal assemblage change.

A first iteration, and perhaps the most direct way of testing for the influence of dispersal, might result in the use of individual-based models (IBMs) with or without dispersal rates that vary among species to generate predictions of recent occupancy changes at a single site or variation in occupancy among sites. Dispersal probability density functions for each species could be parameterized based on either modern traits (e.g. body size) or past rates of range shift (estimated from data in Neotoma). IBMs parameterized this way could then be validated against historical data from Grinnell Resurvey Project (+ancillary museum records) and/or modern data from NEON sites and other ecological surveys conducted in the Sierra Nevada (Kelt et al. 2013). The difference in the goodness of fit to empirical data of IBM models with or without dispersal rates should also provide insight into the importance of dispersal on assemblage change.

A second iteration might result in the use of disequilibrium models to integrate data at single (e.g. Samwell Cave) or multiple sites (e.g. Grinnell Resurvey Project sites and/or NEON sites) together with abiotic environmental data (e.g. climate) over the last ~21 000 years to assess the level to which community-level responses have been at (dis)equilibrium with expectations from static climatic niches. Dispersal lags are thought to be one of the primary, though not necessarily the only, causes of disequilibria with climate. Moreover, paleocommunity ecology data from Faunmap and Neotoma could be used to estimate the direction, magnitude and velocity of past range shifts for individual species and test whether the communities with a higher number of dispersal-limited species show a higher degree of disequilibrium.

Third, once data can 'talk' to each other and speak to the same processes, the main remaining challenge is to account for the different sampling biases inherent to vastly different data-generation processes. Fortunately, different sampling processes are increasingly being accommodated within shared (or joint) modeling frameworks (Fithian et al. 2015, Warton et al. 2015b, Clark et al. 2017, Ovaskainen et al. 2017, Pacifici et al. 2017), and continued progress in this direction, especially with regards to accommodating paleo-ecological data, will be required for broader spatiotemporal integration. Key advances are being made by the PalEON project (<http://sites.nd.edu/paleonproject/>) in this regard (Dawson et al. 2016, Goring et al. 2016). Continued cross-pollination from more distant scientific disciplines (e.g. epidemiology, Bliznyuk et al. 2014) will also be necessary.

Finally, although we believe that a narrowed, sequential focus on one or a subset of processes may facilitate spatio-temporal integration, we do not propose that integrative studies cease to attempt modeling all processes concurrently. Our review also highlights the potential for individual-based and related classes of models to capture a large range of data types relevant to explaining all four community assembly processes (Fig. 2). For example, a recent study was surprisingly successful at reproducing the eco-evolutionary history of South American birds over the past 800 000 years from first principles, using a mechanistic model that accounted for dispersal, selection (both biotic and abiotic) and speciation (Rangel et al. 2018). Additionally, a recent empirical synthesis of global biodiversity based on first principles, and incorporating all of the fundamental processes in Vellend's framework (Vellend 2010, 2016), provides strikingly accurate predictions of latitudinal gradients in species richness for many coastal and terrestrial groups (Worm and Tittensor 2018). Crucially, such mechanistic approaches allow detailed comparisons between empirical and predicted patterns. Data-model comparisons like these represent a valuable additional tool for elucidating community processes in space and time, which transcends current disciplinary boundaries. Moreover, they could help identify key areas of uncertainty where additional data collection could be particularly helpful.

Overall, our hope is that a stronger focus on examining processes across larger scales of space and time may enable greater integration among the theory, data and methods necessary for understanding ecological change across large spatial, temporal and taxonomic scales, at a time where a robust understanding of ecological change is more crucial than ever to tackle today's most pressing environmental challenges.

# References

Abatzoglou, J. T. et al. 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. – Sci. Data 5: 1–12.

Alonso, D. et al. 2006. The merits of neutral theory. – Trends Ecol. Evol. 21: 451–457.

Alroy, J. et al. 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. – Proc. Natl Acad. Sci. USA 98: 6261–6266.

Anderson-Teixeira, K. J. et al. 2015. CTFS-ForestGEO: a worldwide network monitoring forests in an era of global change. – Global Change Biol. 21: 528–549.

Bachelot, B. et al. 2018. Associations among arbuscular mycorrhizal fungi and seedlings are predicted to change with tree successional status. – Ecology 99: 607–620.

Bailey, L. L. et al. 2014. Advances and applications of occupancy models. – Methods Ecol. Evol. 5: 1269–1279.

Bálint, M. et al. 2018. Environmental DNA time series in ecology. – Trends Ecol. Evol. 33: 945–957.

Ballantyne, G. et al. 2017. Pollinator importance networks illustrate the crucial value of bees in a highly speciose plant community. – Sci. Rep. 7: 1–13.

Barnosky, A. D. et al. 2011. Has the Earth's sixth mass extinction already arrived? – Nature 471: 51–57.

Barnosky, A. D. et al. 2017. Merging paleobiology with conservation biology to guide the future of terrestrial ecosystems. – Science 355: 594–606.

Beck, J. et al. 2012. What's on the horizon for macroecology? – Ecography 35: 673–683.

Beckman, N. G. and Rogers, H. S. 2013. Consequences of seed dispersal for plant recruitment in tropical forests: interactions within the seedscape. – Biotropica 45: 666–681.

Belmaker, J. and Jetz, W. 2015. Relative roles of ecological and energetic constraints, diversification rates and region history on global species richness gradients. – Ecol. Lett. 18: 563–571.

Blaauw, M. and Christeny, J. A. 2011. Flexible paleoclimate age-depth models using an autoregressive gamma process. – Bayesian Anal. 6: 457–474.

Bliznyuk, N. et al. 2014. Nonlinear predictive latent process models for integrating spatio–temporal exposure data multiple sources. – Ann. Appl. Stat. 8: 1538–1560.

Blois, J. L. et al. 2010. Small mammal diversity loss in response to late-Pleistocene climatic change. – Nature 465: 771–774.

Blois, J. L. et al. 2013. Climate change and the past, present and future of biotic interactions. – Science 341: 499–504.

Blois, J. L. et al. 2014. A framework for evaluating the influence of climate, dispersal limitation and biotic interactions using fossil pollen associations across the late Quaternary. – Ecography 37: 1095–1108.

Blonder, B. et al. 2015. Linking environmental filtering and disequilibrium to biogeography with a community climate framework. – Ecology 96: 972–985.

Blonder, B. et al. 2017. Predictability in community dynamics. – Ecol. Lett. 20: 293–306.

Bloom, T. D. S. et al. 2018. Why georeferencing matters: introducing a practical protocol to prepare species occurrence records for spatial analysis. – Ecol. Evol. 8: 765–777.

Brereton, T. M. et al. 2017. United Kingdom butterfly monitoring scheme report for 2016. – Centre for Ecology and Hydrology, Wallingford, UK, and Butterfly Conservation, Wareham, UK.

Brewer, S. et al. 2012. Paleoecoinformatics: applying geohistorical data to ecological questions. – Trends Ecol. Evol. 27: 104–112.

Bronk Ramsey, C. 2008. Radiocarbon dating: revolutions in understanding. – Archaeometry 50: 249–275.

Bronk Ramsey, C. 2009. Bayesian analysis of radiocarbon dates. – Radiocarbon 51: 337–360.

Bronk Ramsey, C. et al. 2004. Towards high-precision AMS: progress and limitations. – Radiocarbon 46: 17–24.

Brown, J. H. 2014. Why are there so many species in the tropics? – J. Biogeogr. 41: 8–22.

Brown, A. M. et al. 2014. The fourth-corner solution – using predictive models to understand how species traits interact with the environment. – Methods Ecol. Evol. 5: 344–352.

BTO et al. 2016. The breeding bird survey 2016. – British Trust for Ornithology, Thetford, UK.

Buck, C. E. et al. 1999. BCal: an on-line Bayesian radiocarbon calibration tool. – Internet Archaeol.

Cadotte, M. W. and Davies, J. T. 2016. Phylogenies in ecology: a guide to concepts and methods. – Princeton Univ. Press.

Cadotte, M. W. et al. 2011. Beyond species: functional diversity and the maintenance of ecological processes and services. – J. Appl. Ecol. 48: 1079–1087.

Cadotte, M. et al. 2013. The ecology of differences: assessing community assembly with trait and evolutionary distances. – Ecol. Lett. 16: 1234–1244.

Carlucci, M. B. et al. 2017. Phylogenetic composition and structure of tree communities shed light on historical processes influencing tropical rainforest diversity. – Ecography 40: 521–530.

Cavender-Bares, J. et al. 2009. The merging of community ecology and phylogenetic biology. – Ecol. Lett. 12: 693–715.

Chan, Y. L. et al. 2014. Detecting concerted demographic response across community assemblages using hierarchical approximate Bayesian computation. – Mol. Biol. Evol. 31: 2501–2515.

Chave, J. 2004. Neutral theory and community ecology. – Ecol. Lett. 7: 241–253.

Chesson, P. 2009. Scale transition theory with special reference to species coexistence in a variable environment. – J. Biol. Dynam. 3: 149–163.

Chesson, P. 2017. AEDT: a new concept for ecological dynamics in the ever-changing world. – PLoS Biol. 15: 1–13.

Clark, J. S. 1998. Why trees migrate so fast: confronting theory with dispersal biology and the paleorecord. – Am. Nat. 152: 204–24.

Clark, J. S. et al. 1998. Reid's paradox of plant rapid migration dispersal theory and interpretation of paleoecological records. – Bioscience 48: 13–24.

Clark, J. S. et al. 2003. Estimating population spread: what can we forecast and how well? – Ecology 84: 1979–1988.

Clark, J. S. et al. 2017. Generalized joint attribute modeling for biodiversity analysis: median-zero, multivariate, multifarious data. – Ecol. Monogr. 87: 34–56.

Coelho, M. T. P. and Rangel, T. F. 2018. Neutral community dynamics and the evolution of species interactions. – Am. Nat. 191: 421–434.

Condamine, F. L. et al. 2013. Macroevolutionary perspectives to environmental change. – Ecol. Lett. 16: 72–85.

Cornwell, W. K. et al. 2006. A trait-based test for habitat filtering: convex hull volume. – Ecology 87: 1465–1471.

Correa-Metrio, A. et al. 2014. Environmental determinism and neutrality in vegetation at millennial time scales. – J. Veg. Sci. 25: 627–635.

Crawley, M. et al. 2005. Determinants of species richness in the park grass experiment. – Am. Nat. 165: 179–191.

Currie, D. J. et al. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. – Ecol. Lett. 7: 1121–1134.

Davies, T. J. et al. 2011. Neutral biodiversity theory can explain the imbalance of phylogenetic trees but not the tempo of their diversification. – Evolution 65: 1841–1850.

Dawson, A. et al. 2016. Quantifying pollen-vegetation relationships to reconstruct ancient forests using 19th-century forest composition and pollen data. – Q. Sci. Rev. 137: 156–175.

DeAngelis, D. L. and Grimm, V. 2014. Individual-based models in ecology after four decades. – F1000Prime Rep. in press.

Deiner, K. et al. 2017. Environmental DNA metabarcoding: transforming how we survey animal and plant communities. – Mol. Ecol. 26: 5872–5895.

Desjardins-Proulx, P. and Gravel, D. 2011. How likely is speciation in neutral ecology? – Am. Nat. 179: 137–144.

Desjardins-Proulx, P. and Gravel, D. 2012. A complex speciation-richness relationship in a simple neutral model. – Ecol. Evol. 2: 1781–1790.

Dick, C. W. et al. 2007. Extreme long-distance dispersal of the lowland tropical rainforest tree Ceiba pentandra L. (Malvaceae) in Africa and the Neotropics. – Mol. Ecol. 16: 3039–3049.

Domingo, M. S. et al. 2012. Resource partitioning among top predators in a Miocene food web resource partitioning among top predators in a Miocene food web. – Proc. R. Soc. B 280: 2012138.

Dorazio, R. M. 2014. Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. – Global Ecol. Biogeogr. 23: 1472–1484.

Dorazio, R. M. et al. 2010. Models for inference in dynamic metacommunity systems. – Ecology 91: 2466–2475.

Dray, S. and Legendre, P. 2008. Testing the species traits–environment relationships: the fourth-corner problem revisited. – Ecology 89: 3400–3412.

Dunson, W. A. and Travis, J. 1991. The role of abiotic factors in community organization. – Am. Nat. 138: 1067–1091.

Edwards, J. L. 2004. Research and societal benefits of the global biodiversity information facility. – Bioscience 54: 486.

Ellwood, E. R. et al. 2015. Accelerating the digitization of biodiversity research specimens through online public participation. – Bioscience 65: 383–396.

Enquist, B. J. et al. 2009. The Botanical Information and Ecology Network (BIEN): cyberinfrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity. – iPlant Collab.

Enquist, B. J. et al. 2015. Scaling from traits to ecosystems: developing a general trait driver theory via integrating trait-based and metabolic scaling theories. – Elsevier.

Eronen, J. T. et al. 2010. Ecometrics: the traits that bind the past and present together. – Integr. Zool. 5: 88–101.

Etienne, R. S. et al. 2012. Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. – Proc. R. Soc. B 279: 1300–1309.

Evans, S. et al. 2017. Effects of dispersal and selection on stochastic assembly in microbial communities. – ISME J. 11: 176–185.

Ezard, T. H. G. et al. 2011. Interplay between changing climate and species' ecology drives macroevolutionary dynamics. – Science 332: 349–351.

Ezard, T. H. G. et al. 2012. The meaning of birth and death (in macroevolutionary birth-death models). – Biol. Lett. 8: 139–142.

FAUNMAP Working Group 1994. FAUNMAP: a database documenting late Quaternary distributions of mammal species in the United States. – Illinois State Museum Sci. Pap. 25: 1–690.

Feranec, R. S. et al. 2007. Radiocarbon dates from the Pleistocene fossil deposits of Samwel Cave, Shasta County, California, USA. – Radiocarbon 49: 117–121.

Fick, S. E. and Hijmans, R. J. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. – Int. J. Climatol. 37: 4302–4315.

Fine, P. 2015. Ecological and evolutionary drivers of geographic variation in species diversity. – Annu. Rev. Ecol. Evol. Syst. 46: 369–392.

Fithian, W. et al. 2015. Bias correction in species distribution models: pooling survey and collection data for multiple species. – Methods Ecol. Evol. 6: 424–438.

Foote, M. 2000. Origination and extinction components of taxonomic diversity: general problems. – Paleobiology 26: 74–102.

Fordham, D. A. et al. 2017. PaleoView: a tool for generating continuous climate projections spanning the last 21 000 years at regional and global scales. – Ecography 40: 1348–1358.

Fritz, S. A. et al. 2013. Diversity in time and space: wanted dead and alive. – Trends Ecol. Evol. 28: 509–516.

Fuller, M. M. et al. 2005. Effects of predation and variation in species relative abundance on the parameters of neutral models. – Comm. Ecol. 6: 229–240.

Funk, J. L. et al. 2017. Revisiting the Holy Grail: using plant functional traits to understand ecological processes. – Biol. Rev. 92: 1156–1173.

Furlong, E. 1904. An account of the preliminary excavations in a recently explored quaternary cave in Shasta County, California. – Science 20: 53–55.

Furlong, E. 1906. The exploration of Samwel Cave. – Am. J. Sci. 172: 235–247.

Galic, N. et al. 2018. When things don't add up: quantifying impacts of multiple stressors from individual metabolism to ecosystem processing. – Ecol. Lett. 21: 568–577.

Gaston, K. J. 2000. Global patterns in biodiversity. - Nature 405: 220–227.

Gaston, K. J. and Blackburn, T. M. 2000. Pattern and process in macroecology. – Blackwell.

Gianuca, A. T. et al. 2017. Integrating trait and phylogenetic distances to assess scale-dependent community assembly processes. – Ecography 40: 742–752.

Godoy, O. et al. 2018. Towards the integration of niche and network theories. – Trends Ecol. Evol. 33: 287–300.

Gorelick, N. et al. 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. – Remote Sens. Environ. 202: 18–27.

Goring, S. J. et al. 2016. Novel and lost forests in the upper Midwestern United States, from new estimates of settlement-era composition, stem density and biomass. – PLoS One 11: e0151935.

Graham, L. J. et al. 2017. Future geographic patterns of novel and disappearing assemblages across three dimensions of diversity: a case study with Ecuadorian hummingbirds. – Divers. Distrib. 23: 944–954.

Gravel, D. et al. 2016. Stability and complexity in model meta-ecosystems. – Nat. Commun. 7: 12457.

Grimm, V. et al. 2006. A standard protocol for describing individual-based and agent-based models. – Ecol. Model. 198: 115–126.

Gross, N. et al. 2009. Linking individual response to biotic interactions with community structure: a trait-based framework. – Funct. Ecol. 23: 1167–1178.

Guillera-Arroita, G. 2017. Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. – Ecography 40: 281–295.

Guralnick, R. 2006. The legacy of past climate and landscape change on species'current experienced climate and elevation ranges across latitude: a multispecies study utilizing mammals in western North America. – Global Ecol. Biogeogr. 15: 505–518.

Guralnick, R. P. et al. 2007. Towards a collaborative, global infrastructure for biodiversity assessment. – Ecol. Lett. 10: 663–672.

Haase, P. et al. 2018. The next generation of site-based long-term ecological monitoring: linking essential biodiversity variables and ecosystem integrity. – Sci. Total Environ. 613–614: 1376–1384.

Hanson, C. A. et al. 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. – Nat. Rev. Microbiol. 10: 497–506.

Hardisty, A. and Roberts, D. 2013. A decadal view of biodiversity informatics: challenges and priorities. – BMC Ecol. 13: 1–23.

Harte, J. 2011. Maximum entropy and ecology: a theory of abundance, distribution and energetics. – Oxford Univ. Press.

Haslett, J. and Parnell, A. 2008. A simple monotone process with application to radiocarbon-dated depth chronologies. – J. R. Stat. Soc. Ser. C Appl. Stat. 57: 399–418.

Heintzman, P. D. et al. 2017. A new genus of horse from pleistocene North America. – Elife 6: 1–43.

Heleno, R. et al. 2014. Ecological networks: delving into the architecture of biodiversity. – Biol. Lett. 10: 20131000.

Hickerson, M. J. et al. 2010. Phylogeography's past, present and future: 10 years after Avise, 2000. – Mol. Phylogenet. Evol. 54: 291–301.

Hickler, T. et al. 2009. An ecosystem model-based estimate of changes in water availability differs from water proxies that are

commonly used in species distribution models. – Global Ecol. Biogeogr. 18: 304–313.

Hobbie, J. et al. 2003. The US long term ecological research program. – Bioscience 53: 21–32.

Holland, S. M. 2018. Diversity and tectonics: predictions from neutral theory. – Paleobiology 44: 219–236.

Holt, G. and Chesson, P. 2016. Scale-dependent community theory for streams and other linear habitats. – Am. Nat. 188: E59–E73.

Holt, R. D. and Gaines, M. S. 1992. Analysis of adaptation in heterogeneous landscapes: implications for the evolution of fundamental niches. – Evol. Ecol. 6: 433–447.

Horvát, S. et al. 2010. A spatially explicit model for tropical tree diversity patterns. – J. Theor. Biol. 265: 517–523.

Hu, X.-S. et al. 2006. Neutral theory in macroecology and population genetics. – Oikos 113: 548–555.

Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. – Princeton Univ. Press.

Hubert, N. et al. 2015. Metacommunity speciation models and their implications for diversification theory. – Ecol. Lett. 18: 864–881.

Iknayan, K. J. et al. 2014. Detecting diversity: emerging methods to estimate species diversity. – Trends Ecol. Evol. 29: 97–106.

Isaac, N. J. B. et al. 2014. Statistics for citizen science: extracting signals of change from noisy ecological data. – Methods Ecol. Evol. 5: 1052–1060.

IUCN 2019. The IUCN Red List of threatened species. Ver. 2019-1. – <http://www.iucnredlist.org>.

Jablonski, D. et al. 2006. Out of the tropics: evolutionary diversity gradient. – Science 314: 102–106.

Jablonski, D. et al. 2017. Shaping the latitudinal diversity gradient: new perspectives from a synthesis of paleobiology and biogeography. – Am. Nat. 189: 1–12.

Jabot, F. 2010. A stochastic dispersal-limited trait-based model of community dynamics. – J. Theor. Biol. 262: 650–661.

Jackson, S. T. 2012. Representation of flora and vegetation in Quaternary fossil assemblages: known and unknown knowns and unknowns. – Q. Sci. Rev. 49: 1–15.

Jackson, S. T. and Blois, J. L. 2015. Community ecology in a changing environment: perspectives from the Quaternary. – Proc. Natl Acad. Sci. USA 112: 4915–4921.

Jackson, S. T. et al. 2005. A 40 000-year woodrat-midden record of vegetational and biogeographical dynamics in north-eastern Utah, USA. – J. Biogeogr. 32: 1085–1106.

Janzen, T. et al. 2015. A sampling formula for ecological communities with multiple dispersal syndromes. – J. Theor. Biol. 374: 94–106.

Jones, K. E. et al. 2009. PanTHERIA: a species-level database of life history, ecology and geography of extant and recently extinct mammals. – Ecology 90: 2648–2648.

Jønsson, K. A. et al. 2016. Tracking animal dispersal: from individual movement to community assembly and global range dynamics. – Trends Ecol. Evol. 31: 204–214.

Karger, D. N. et al. 2017. Climatologies at high resolution for the earth's land surface areas. – Sci. Data 4: 1–20.

Kattge, J. et al. 2011. TRY – a global database of plant traits. – Global Change Biol. 17: 2905–2935.

Kearney, M. R. et al. 2014. Microclim: global estimates of hourly microclimate based on long-term monthly climate averages. – Sci. Data 1: 1–9.

Keller, M. et al. 2008. A continental strategy for the National Ecological Observatory Network. – Front. Ecol. Environ. 6: 282–284.

Kelly, M. et al. 2016. Rescuing and sharing historical vegetation data for ecological analysis: the California vegetation type mapping project. – Biodivers. Inform. 11: 40–62.

Kelly, M. et al. 2017. Geospatial informatics key to recovering and sharing historical ecological data for modern use. – Photogramm. Eng. Remote Sens. 83: 779–786.

Kelt, D. et al. 2013. Small mammals exhibit limited spatiotemporal structure in Sierra Nevada forests. – J. Mammal. 94: 1197–1213.

Kissling, W. D. et al. 2016. Historical colonization and dispersal limitation supplement climate and topography in shaping species richness of African lizards (Reptilia: Agaminae). – Sci. Rep. 6: 1–14.

Kissling, W. D. et al. 2017. Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. – Biol. Rev. 93: 600–625.

Kriticos, D. J. et al. 2012. CliMond: global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. – Methods Ecol. Evol. 3: 53–64.

Lam, A. R. et al. 2018. Dispersal in the Ordovician: speciation patterns and paleobiogeographic analyses of brachiopods and trilobites. – Palaeogeogr. Palaeoclimatol. Palaeoecol. 489: 147–165.

Lamichhaney, S. et al. 2017. Rapid hybrid speciation in Darwin's finches. – Science 4593: 224–228.

Lamsdell, J. C. et al. 2017. Phylogenetic paleoecology: tree-thinking and ecology in deep time. – Trends Ecol. Evol. 32: 452–463.

Lasky, J. R. et al. 2014. Trait-mediated assembly processes predict successional changes in community diversity of tropical forests. – Proc. Natl Acad. Sci. USA 111: 5616–5621.

Latombe, G. et al. 2015. Beyond the continuum: a multi-dimensional phase space for neutral–niche community assembly. – Proc. R. Soc. B 282: 20152417.

Laughlin, D. C. and Messier, J. 2015. Fitness of multidimensional phenotypes in dynamic adaptive landscapes. – Trends Ecol. Evol. 30: 487–496.

Laughlin, D. C. et al. 2012. A predictive model of community assembly that incorporates intraspecific trait variation. – Ecol. Lett. 15: 1291–1299.

Lawton, J. H. 1999. Are there general laws in ecology? – Oikos 84: 177–192.

Leibold, M. A. and Chase, J. M. 2018. Metacommunity ecology. – Princeton Univ. Press.

Leibold, M. a. et al. 2004. The metacommunity concept: a framework for multi-scale community ecology. – Ecol. Lett. 7: 601–613.

Lessard, J.-P. et al. 2016. Process-based species pools reveal the hidden signature of biotic interactions amid the influence of temperature filtering. – Am. Nat. 187: 75–88.

Levin, S. A. 2012. Towards the marriage of theory and data. – Interface Focus 2: 141–143.

Levine, J. M. and Murrell, D. J. 2003. The community-level consequences of seed dispersal patterns. – Annu. Rev. Ecol. Evol. Syst. 34: 549–574.

Lindenmayer, D. B. et al. 2012. Value of long-term ecological studies. – Austral Ecol. 37: 745–757.

Liow, L. H. 2013. Simultaneous estimation of occupancy and detection probabilities: an illustration using Cincinnatian brachiopods. – Paleobiology 39: 193–213.

Lischke, H. et al. 2006. TreeMig: a forest-landscape model for simulating spatio-temporal patterns from stand to landscape scale. – Ecol. Model. 199: 409–420.

Loarie, S. R. et al. 2009. The velocity of climate change. – Nature 462: 1052–1055.

Loranger, J. et al. 2016. Recasting the dynamic equilibrium model through a functional lens: the interplay of trait-based community assembly and climate. – J. Ecol. 104: 781–791.

Lososová, Z. et al. 2015. Phylogenetic structure of plant species pools reflects habitat age on the geological time scale. – J. Veg. Sci. 26: 1080–1089.

Luo, Y. et al. 2011. Ecological forecasting and data assimilation in a data-rech era. – Ecol. Appl. 21: 1429–1442.

Lyons, S. K. et al. 2016. Holocene shifts in the assembly of plant and animal communities implicate human impacts. – Nature 529: 80–83.

Maguire, K. C. et al. 2015. Modeling species and community responses to past, present and future episodes of climatic and ecological change. – Annu. Rev. Ecol. Evol. Syst. 46: 343–368.

Magurran, A. E. 2007. Species abundance distributions over time. – Ecol. Lett. 10: 347–354.

Magurran, A. E. et al. 2010. Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. – Trends Ecol. Evol. 25: 574–582.

Marin, J. and Hedges, S. B. 2016. Time best explains global variation in species richness of amphibians, birds and mammals. – J. Biogeogr. 43: 1069–1079.

Marin, J. et al. 2018. Evolutionary time drives global tetrapod diversity. – Proc. R. Soc. B 285: 5–12.

Marquet, P. a. et al. 2014. On theory in ecology. – Bioscience 64: 701–710.

Marske, K. A. et al. 2013. Phylogeography: spanning the ecology–evolution continuum. – Ecography 36: 1169–1181.

Mason, N. W. H. et al. 2013. A guide for using functional diversity indices to reveal changes in assembly processes along ecological gradients. – J. Veg. Sci. 24: 794–806.

May, R. M. 2004. Uses and abuses of mathematics in biology. – Science 303: 790–793.

Mayfield, M. M. and Levine, J. M. 2010. Opposing effects of competitive exclusion on the phylogenetic structure of communities. – Ecol. Lett. 13: 1085–1093.

Mayfield, M. M. et al. 2010. What does species richness tell us about functional trait diversity? Predictions and evidence for responses of species and functional trait diversity to land-use change. – Global Ecol. Biogeogr. 19: 423–431.

Mayhew, P. J. et al. 2008. A long-term association between global temperature and biodiversity, origination and extinction in the fossil record. – Proc. R. Soc. B 275: 47–53.

Mccarthy, H. R. et al. 2012. Integrating empirical-modeling approaches to improve understanding of terrestrial ecology processes. – New Phytol. 195: 523–525.

McGill, B. J. 2010. Towards a unification of unified theories of biodiversity. – Ecol. Lett. 13: 627–642.

McGill, B. J. 2019. The what, how and why of doing macroecology. – Global Ecol. Biogeogr. 28: 6–17.

McGill, B. J. et al. 2005. Community inertia of Quaternary small mammal assemblages in North America. – Proc. Natl Acad. Sci. USA 102: 16701–16706.

McGill, B. J. et al. 2006a. Empirical evaluation of neutral theory. – Ecology 87: 1411–1423.

McGill, B. J. et al. 2006b. Rebuilding community ecology from functional traits. – Trends Ecol. Evol. 21: 178–185.

McGill, B. J. et al. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. – Ecol. Lett. 10: 995–1015.

McLachlan, J. S. and Clark, J. S. 2004. Reconstructing historical ranges with fossil data at continental scales. – For. Ecol. Manage. 197: 139–147.

McPeek, M. A. and Brown, J. M. 2000. Building a regional species pool: diversification of the *Enallagma* damselflies in eastern North America. – Ecology 81: 904–920.

Melbourne, B. A. and Chesson, P. 2005. Scaling up population dynamics: integrating theory and data. – Oecologia 145: 179–187.

Messier, J. et al. 2010. How do traits vary across ecological scales? A case for trait-based ecology. – Ecol. Lett. 13: 838–848.

Metcalf, J. L. et al. 2014. Integrating multiple lines of evidence into historical biogeography hypothesis testing: a *Bison bison* case study. – Proc. R. Soc. B 281: 20132782–20132782.

Meyer, C. et al. 2015. Global priorities for an effective information basis of biodiversity distributions. – Nat. Commun. 6: 1–8.

Meyer, C. et al. 2016. Multidimensional biases, gaps and uncertainties in global plant occurrence information. – Ecol. Lett.: 992–1006.

Michener, W. K. and Jones, M. B. 2012. Ecoinformatics: supporting ecology as a data-intensive science. – Trends Ecol. Evol. 27: 88–93.

Mittelbach, G. G. and Schemske, D. W. 2015. Ecological and evolutionary perspectives on community assembly. – Trends Ecol. Evol. 30: 241–247.

Montoya, D. et al. 2015. Functional group diversity increases with modularity in complex food webs. – Nat. Commun. 6: 1–9.

Morlon, H. et al. 2011. Reconciling molecular phylogenies with the fossil record. – Proc. Natl Acad. Sci. USA 108: 16327–16332.

Nathan, R. 2006. Long-distance dispersal of plants. – Science 313: 786–8.

Nathan, R. et al. 2011. Spread of North American wind-dispersed trees in future environments. – Ecol. Lett. 14: 211–219.

Nathan, R. et al. 2012. Dispersal kernels: review. – In: Clobert, J, et al. (eds), Dispersal ecology and evolution. Oxford Univ. Press.

Nemergut, D. R. et al. 2013. Patterns and processes of microbial community assembly. – Microbiol. Mol. Biol. Rev. 77: 342–356.

Norberg, A. et al. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. – Ecol. Monogr. in press.

Norberg, J. J. et al. 2001. Phenotypic diversity and ecosystem functioning in changing environments: a theoretical framework. – Proc. Natl Acad. Sci. USA 98: 11376–11381.

Oliveira, B. F. et al. 2016. Species and functional diversity accumulate differently in mammals. – Global Ecol. Biogeogr. 25: 1119–1130.

Ordonez, A. and Williams, J. W. 2013. Climatic and biotic velocities for woody taxa distributions over the last 16 000 years in eastern North America. – Ecol. Lett. 16: 773–781.

Ovaskainen, O. et al. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. – Ecol. Lett. 20: 561–576.

Pacifici, K. et al. 2017. Integrating multiple data sources in species distribution modeling: a framework for data fusion. – Ecology 98: 840–850.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. – Nature 401: 877–884.

Parr, C. S. et al. 2012. Evolutionary informatics: unifying knowledge about the diversity of life. – Trends Ecol. Evol. 27: 94–103.

Pásztor, L. et al. 2016. Theory-based ecology. - Oxford Univ. Press.

Patzkowsky, M. E. 2017. Origin and evolution of regional biotas: a deep-time perspective. – Annu. Rev. Earth Planet. Sci. 45: 471–495.

Pearse, W. D. et al. 2018. Building up biogeography: pattern to process. – J. Biogeogr. 45: 1223–1230.

Pedersen, M. W. et al. 2015. Ancient and modern environmental DNA. – Phil. Trans. R. Soc. B 370: 20130383.

Pennell, M. W. and Harmon, L. J. 2013. An integrative view of phylogenetic comparative methods: connections to population genetics, community ecology and paleobiology. – Ann. N. Y. Acad. Sci. 1289: 90–105.

Pennington, R. T. and Dick, C. W. 2004. The role of immigrants in the assembly of the South American rainforest tree flora. – Phil. Trans. R. Soc. B 359: 1611–1622.

Pigot, A. L. and Etienne, R. S. 2015. A new dynamic null model for phylogenetic community structure. – Ecol. Lett. 18: 153–163.

Pilosof, S. et al. 2017. The multilayer nature of ecological networks. – Nat. Ecol. Evol. 1: 0101.

Pirzamanbein, B. et al. 2014. Creating spatially continuous maps of past land cover from point estimates: a new statistical approach applied to pollen data. – Ecol. Complex. 20: 127–141.

Poelen, J. H. et al. 2014. Global biotic interactions: an open infrastructure to share and analyze species-interaction datasets. – Ecol. Inform. 24: 148–159.

Poisot, T. et al. 2015. Beyond species: why ecological interaction networks vary through space and time. – Oikos 124: 243–251.

Pollock, L. J. et al. 2014. Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). – Methods Ecol. Evol. 5: 397–406.

Pompanon, F. et al. 2012. Who is eating what: diet assessment using next generation sequencing. – Mol. Ecol. 21: 1931–1950.

Porter, J. H. et al. 2012. Staying afloat in the sensor data deluge. – Trends Ecol. Evol. 27: 121–129.

Price, S. A. and Schmitz, L. 2016. A promising future for integrative biodiversity research: an increased role of scale-dependency and functional biology. – Phil. Trans. R. Soc. B 371: 20150228.

Pyle, R. L. 2016. Towards a global names architecture: the future of indexing scientific names. – Zookeys 2016: 261–281.

Pyle, R. L. and Michel, E. 2008. ZooBank: developing a nomenclatural tool for unifying 250 years of biological information. – Zootaxa 50: 39–50.

Quental, T. B. and Marshall, C. R. 2010. Diversity dynamics: molecular phylogenies need the fossil record. – Trends Ecol. Evol. 25: 435–441.

Rangel, T. F. et al. 2018. Modeling the ecology and evolution of biodiversity: biogeographical cradles, museums and graves. – Science 361: eaar5452.

Rapacciuolo, G. et al. 2014a. Beyond a warming fingerprint: individualistic biogeographic responses to heterogeneous climate change in California. – Global Change Biol. 20: 2841–2855.

Rapacciuolo, G. et al. 2014b. Temporal validation plots: quantifying how well correlative species distribution models predict species' range changes over time. – Methods Ecol. Evol. 5: 407–420.

Rapacciuolo, G. et al. 2017. Detecting long-term occupancy changes in Californian odonates from natural history and citizen science records. – Biodivers. Conserv. 26: 2933–2949.

Reimer, P. J. et al. 2009. INTCAL 09 and MARINE09 aadiocarbon age calibration curves, 0–50 000 years Cal BP. – Radiocarbon 51: 1111–1150.

Reimer, P. J. et al. 2013. Intcal13 and Marine13 radiocarbon age calibration curves 0–50 000 years Cal Bp. – Radiocarbon 55: 1869–1887.

Robledo-Arnuncio, J. J. et al. 2014. Space, time and complexity in plant dispersal ecology. – Mov. Ecol. 2: 1–17.

Rocchini, D. et al. 2011. Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. – Prog. Phys. Geogr. 35: 211–226.

Rosindell, J. et al. 2011. The unified neutral theory of biodiversity and biogeography at age ten. – Trends Ecol. Evol. 26: 340–348.

Rosindell, J. et al. 2012. The case for ecological neutral theory. – Trends Ecol. Evol. 27: 203–208.

Rowe, R. J. and Terry, R. C. 2014. Small mammal responses to environmental change: integrating past and present dynamics. – J. Mammal. 95: 1157–1174.

Rowe, K. C. et al. 2014. Spatially heterogeneous impact of climate change on small mammals of montane California. – Proc. R. Soc. B 282: 20141857.

Roy, D. B. and Sparks, T. H. 2000. Phenology of British butterflies and climate change. – Global Change Biol. 6: 407–416.

Royle, J. A. et al. 2007. Hierarchical spatial models of abundance and occurance from imperfect survey data. – Ecol. Monogr. 77: 465–481.

Ruete, A. 2015. Displaying bias in sampling effort of data accessed from biodiversity databases using ignorance maps. – Biodivers. Data J. 3: e5361.

Sauer, J. R. et al. 2017. The first 50 years of the North American breeding bird survey. – Condor 119: 576–593.

Savage, V. M. et al. 2007. A general multi-trait-based framework for studying the effects of biodiversity on ecosystem functioning. – J. Theor. Biol. 247: 213–229.

Scherer, C. et al. 2016. Merging trait-based and individual-based modelling: an animal functional type approach to explore the responses of birds to climatic and land use changes in semi-arid African savannas. – Ecol. Model. 326: 75–89.

Schliep, E. M. et al. 2018. Joint species distribution modelling for spatio-temporal occurrence and ordinal abundance data. – Global Ecol. Biogeogr. 27: 142–155.

Scholes, R. J. et al. 2008. Ecology. Toward a global biodiversity observing system. – Science 321: 1044–1045.

Sclafani, J. A. and Holland, S. M. 2013. The species–area relationship in the late ordovician: a test using neutral theory. – Diversity 5: 240–262.

Seddon, A. W. R. et al. 2014. Looking forward through the past: identification of 50 priority research questions in palaeoecology. – J. Ecol. 102: 256–267.

Sepkoski, J. J. 2007. Rates of speciation in the fossil record. – Differentiation 353: 315–326.

Shaffer, H. B. et al. 1998. The role of natural history collections in documenting species declines. – Trends Ecol. Evol. 13: 27–30.

Shipley, B. 2010. Community assembly, natural selection and maximum entropy models. – Oikos 119: 604–609.

Shipley, B. et al. 2006. From plant traits to plant communities: a statistical mechanistic approach to biodiversity. – Science 314: 812–814.

Silvertown, J. et al. 2006. Phylogeny and the hierarchical organisation of plant diversity. – Ecology 87: S39–S49.

Simberloff, D. 2004. Community ecology: is it time to move on? – Am. Nat. 163: 787–799.

Smith, W. B. 2002. Forest inventory and analysis: a national inventory and monitoring program. – Environ. Pollut. 116: S233–S242.

Smith, F. A. et al. 2016. Unraveling the consequences of the terminal Pleistocene megafauna extinction on mammal community assembly. – Ecography 39: 223–239.

Smith, F. A. et al. 2018. Body size downgrading ofmammals over the late Quaternary. – Science 313: 310–313.

Stadler, T. and Bokma, F. 2013. Estimating speciation and extinction rates for phylogenies of higher taxa. – Syst. Biol. 62: 220–230.

Stegen, J. C. et al. 2015. Estimating and mapping ecological processes influencing microbial community assembly. – Front. Microbiol. 6: 1–15.

Stropp, J. et al. 2016. Mapping ignorance: 300 years of collecting flowering plants in Africa. – Global Ecol Biogeogr. 25: 1085–1096.

Suding, K. N. et al. 2008. Scaling environmental change through the community-level: a trait-based response-and-effect framework for plants. – Global Change Biol. 14: 1125–1140.

Sugita, S. 2007a. Theory of quantitative reconstruction of vegetation II: all you need is LOVE. – Holocene 2: 243–257.

Sugita, S. 2007b. Theory of quantitative reconstruction of vegetation I: pollen from large sites REVEALS regional vegetation composition. – Holocene 2: 229–241.

Suttle, K. B. et al. 2007. Species interactions reverse grassland responses to changing climate. – Science 315: 640–642.

Svenning, J.-C. and Skov, F. 2004. Limited filling of the potential range in European tree species. – Ecol. Lett. 7: 565–573.

Svenning, J.-C. and Skov, F. 2007. Could the tree diversity pattern in Europe be generated by postglacial dispersal limitation? – Ecol. Lett. 10: 453–460.

Svenning, J. C. et al. 2008. Glacial refugia of temperate trees in Europe: insights from species distribution modelling. – J. Ecol. 96: 1117–1127.

Svenning, J. et al. 2015. The influence of paleoclimate on present-day patterns in biodiversity and ecosystems. – Annu. Rev. Ecol. Evol. Syst. 46: 551–572.

Swengel, A. B. 1990. Monitoring butterfly populations using the Fourth of July Butterfly Count. – Am. Midl. Nat. 124: 395–406.

Taberlet, P. et al. 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. – Mol. Ecol. 21: 2045–2050.

Taylor-Rodriguez, D. et al. 2017. Joint species distribution modeling: dimension reduction using dirichlet processes. – Bayesian Anal. 12: 939–967.

Terry, R. C. 2010. The dead do not lie: using skeletal remains for rapid assessment of historical small-mammal community baselines. – Proc. R. Soc. B 277: 1193–1201.

Terry, R. C. and Rowe, R. J. 2015. Energy flow and functional compensation in Great Basin small mammals under natural and anthropogenic environmental change. – Proc. Natl Acad. Sci. USA 112: 9656–9661.

Tessarolo, G. et al. 2017. Temporal degradation of data limits biodiversity research. – Ecol. Evol. 7: 6863–6870.

Thorson, J. T. et al. 2016. Joint dynamic species distribution models: a tool for community ordination and spatio-temporal monitoring. – Global Ecol. Biogeogr. 25: 1144–1158.

Tikhonov, G. et al. 2017. Using joint species distribution models for evaluating how species-to-species associations depend on the environmental context. - Methods Ecol. Evol. 8: 443–452.

Tilman, D. et al. 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. – Nature 379: 718–720.

Tingley, M. W. and Beissinger, S. R. 2009. Detecting range shifts from historical species occurrences: new perspectives on old data. – Trends Ecol. Evol. 24: 625–633.

Tingley, M. W. and Beissinger, S. R. 2013. Cryptic loss of montane avian richness and high community turnover over 100 years. – Ecology 94: 598–609.

Tingley, M. W. et al. 2012. The push and pull of climate change causes heterogeneous shifts in avian elevational ranges. – Global Change Biol. 18: 3279–3290.

Toju, H. et al. 2017. Species-rich networks and eco-evolutionary synthesis at the metacommunity level. – Nat. Ecol. Evol. 1: 1–11.

Tomašových, A. and Kidwell, S. M. 2017. Nineteenth-century collapse of a benthic marine ecosystem on the open continental shelf. – Proc. R. Soc. B 284: 20170328.

Trakhtenbrot, A. et al. 2005. The importance of long distance dispersal in biodiversity conservation. – Divers. Distrib. 11: 173–181.

Troudet, J. et al. 2017. Taxonomic bias in biodiversity data and societal preferences. – Sci. Rep. 7: 9132.

Urban, M. C. et al. 2012. On a collision course: competition and dispersal differences create no-analogue communities and cause extinctions during climate change. – Proc. R. Soc. B 279: 2072–2080.

Uriarte, M. et al. 2009. Natural and human disturbance land use as determinants of tropical forest results from a forest simulator dynamics: results from a forest simulation. – Ecol. Monogr. 79: 423–443.

Van Strien, A. J. et al. 2011. Metapopulation dynamics in the butterfly *Hipparchia semele* changed decades before occupancy declined in the Netherlands. – Ecol. Appl. 21: 2510–20.

Van Strien, A. J. et al. 2013. Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. – J. Appl. Ecol. 50: 1450–1458.

Vega, G. C. et al. 2017. Data descriptor: MERRAclim, a high-resolution global dataset of remotely sensed bioclimatic variables for ecological modelling. – Sci. Data 4: 170078.

Vellend, M. 2010. Conceptual synthesis in community ecology. – Q. Rev. Biol. 85: 183–206.

Vellend, M. 2016. The theory of ecological communities. – Princeton Univ. Press.

Vellend, M. et al. 2013. Historical ecology: using unconventional data sources to test for effects of global environmental change. – Am. J. Bot. 100: 1294–1305.

Violle, C. et al. 2014. The emergence and promise of functional biogeography. – Proc. Natl Acad. Sci. USA 111: 13690–13696.

Warton, D. I. et al. 2015a. CATS regression – a model-based approach to studying trait-based community assembly. – Methods Ecol. Evol. 6: 389–398.

Warton, D. I. et al. 2015b. So many variables: joint modeling in community ecology. – Trends Ecol. Evol. 30: 766–779.

Weathers, K. C. et al. 2016. Frontiers in ecosystem ecology from a community perspective: the future is boundless and bright. – Ecosystems 19: 753–770.

Whitmee, S. and Orme, C. D. L. 2012. Predicting dispersal distance in mammals: a trait-based approach. – J. Anim. Ecol. 82: 211–221.

Wiezorek, J. et al. 2012. Darwin core: an evolving community-developed biodiversity data standard. – PLoS One 7: e29715.

Williams, J. E. and Blois, J. L. 2018. Range shifts in response to past and future climate change: can climate velocities and species' dispersal capabilities explain variation in mammalian range shifts? – J. Biogeogr. 45: 2175–2189.

Williams, J. W. et al. 2018. The Neotoma Paleoecology Database, a multiproxy, international, community-curated data resource. – Q. Res. 89: 156–177.

Wood, R. 2015. From revolution to convention: the past, present and future ofradiocarbon dating. – J. Archaeol. Sci. 56: 61–72.

Worm, B. and Tittensor, D. 2018. A theory of global biodiversity. – Princeton Univ. Press.

Yamaura, Y. et al. 2011. Modelling community dynamics based on species-level abundance models from detection/nondetection data. – J. Appl. Ecol. 48: 67–75.

Yeakel, J. D. et al. 2014. Collapse of an ecological network in Ancient Egypt. – Proc. Natl Acad. Sci. USA 111: 14472–14477.

Ytow, N. 2016. Taxonaut: an application software for comparative display of multiple taxonomies with a use case of GBIF Species API. – Biodivers. Data J. 4: e9787.

Zeilinger, A. R. et al. 2017. Museum specimen data reveal emergence of a plant disease may be linked to increases in the insect vector population. – Ecol. Appl. 27: 1827–1837.

Zhou, J. and Ning, D. 2017. Stochastic community assembly: does it matter in microbial ecology? – Microbiol. Mol. Biol. Rev. 81: 1–32.

Zurell, D. et al. 2018. Do joint species distribution models reliably detect interspecific interactions from co-occurrence data in homogenous environments? – Ecography 41: 1–8.