

ECOGRAPHY

Review and synthesis

Inference of biogeographic history by formally integrating distinct lines of evidence: genetic, environmental niche and fossil

Sean Hoban, Andria Dawson, John D. Robinson, Adam B. Smith and Allan E. Strand

S. Hoban (<https://orcid.org/0000-0002-0348-8449>) ✉ (shoban@mortonarb.org), Center for Tree Science, The Morton Arboretum, Lisle, IL 60532, USA. – *A. Dawson*, Dept of General Education, Mount Royal Univ., Calgary, AB, Canada. – *J. D. Robinson*, Dept of Fisheries and Wildlife, Michigan State Univ., East Lansing, MI, USA. – *A. B. Smith*, Center for Conservation and Sustainable Development, Missouri Botanical Garden, Saint Louis, MO, USA. – *A. E. Strand*, Grice Marine Laboratory, Dept of Biology, College of Charleston, Charleston, SC, USA.

Ecography

42: 1–21, 2019

doi: 10.1111/ecog.04327

Subject Editor: Jessica Blois

Editor-in-Chief:

David Nogués-Bravo

Accepted 4 April 2019

A primary focus of historical biogeography is to understand changes in species ranges, abundance and genetic connectivity, and changes in community composition. Traditionally, biogeographic inference has relied on distinct lines of evidence, including DNA sequences, fossils and hindcasted ecological niche models. In this review we propose that the development of integrative modeling approaches that leverage multiple distinct data types from diverse disciplines has the potential to revolutionize the field of biogeography. Although each data type contains information on a distinct aspect of species' biogeographic histories, few studies formally integrate multiple types in analysis. For example, post hoc congruence among analyses based on different data types (e.g. fossils and genetics) is commonly assumed to indicate likely biogeographic histories. Unfortunately, analyses of different data often reach discordant conclusions. Thus, fundamental and unresolved debates continue regarding speed and timing of postglacial migration, location and size of glacial refugia, and degree of long distance dispersal. Formal statistical integration can help address these issues. More specifically, formal integration can leverage all available evidence, account for inherent biases associated with different data types, and quantify data and process uncertainty. Novel, quantitative integration of data and models across fields is now possible due to recent advances in cyberinfrastructure, spatial modeling, online and aggregated ecological databases, data processing and quantitative methods. Our purpose is to make the case for and give examples of rigorous integration of genetic, fossil and environmental/occurrence data for inferring biogeographic history. In particular, we 1) review the need for such a framework; 2) explain common data types and approaches used to infer biogeographic history (and the challenges with each); 3) review state-of-the-art examples of data integration in biogeography; 4) lay out a series of novel, suggested improvements on current methods; and 5) provide an outlook on technical feasibility and future opportunities.

Keywords: approximate Bayesian computation, Bayesian, climate change, paleoecology, range shifts, species distribution models



Glossary

- ABC (approximate Bayesian computation; also see Box 1) – a simulation-based statistical method that approximates the likelihood of a model through comparisons of summary statistics calculated from simulated and observed data.
- Biogeographic history – the study of processes and patterns related to: geographic range shifts, contractions and expansions; demographic fluctuations; locations of refugia; and migration patterns.
- Bottleneck – a drastic reduction in population size that often also results in a loss of genetic diversity; colonization at a continental scale often involves serial bottlenecks.
- Coalescent theory – a body of theory that uses probabilistic models to reconstruct the genealogy of a sample backward in time from the sampled lineages (present) to the most recent common ancestor (past).
- DRM (dynamic range model) – a hierarchical Bayesian approach to estimate parameter values of species' range dynamics, including dispersal and demographic rates dependent on environmental conditions.
- ENM (ecological niche model) – as used here, a correlative model for predicting suitable environment of a species.
- Genetic differentiation – differences in allele frequencies between two populations.
- Haplotype – a segment of DNA inherited from a single parent, often used in phylogeography as representing a distinct lineage.
- PVM (pollen vegetation model) – a mathematical or statistical model used to estimate vegetation from fossil pollen count records.

Introduction

A major goal of ecologists and evolutionary biologists is to document and understand species' historical dynamics in space and time, i.e. where they existed (phylogeography), with what abundance and genetic connectivity (population biology), with what other species (community ecology), and due to what ecological processes. We refer to the study of species' historical dynamics in space as biogeographic history. The science of biogeographic history includes the study of processes and patterns related to geographic range shifts, contractions and expansions, demographic fluctuations and changes in migration rates. The focus of many such investigations, and the focus of this review, is biogeographic history during the Quaternary (such as species' dynamics during or following glacial periods or Holocene temperature fluctuations). Understanding recent biogeographic history provides an opportunity to document pre-anthropogenic ecosystems and assess how species respond to rapid climatic shifts. Investigations at this temporal and spatial scale can draw upon diverse data sources, including paleoclimate models with relatively high spatial (~100 km) and temporal (~10–100 years; Lorenz et al. 2016, Fordham et al. 2017) resolution during

this time period, and extensive, well-curated records of fossils (especially pollen from sediment cores). DNA from living organisms or fossils also holds information about the past at a range of time scales, including population-level demographic processes that occurred within tens to hundreds of generations (Luikart et al. 1998, Dellicour et al. 2014). Processes such as intercontinental dispersal, speciation and human-mediated events, e.g. crop domestication and invasions, are aspects of biogeographic history, but are outside the focus of this paper. Nonetheless, some of the methods we discuss can be used in these areas.

During climate fluctuations of the Quaternary, species experienced dramatic shifts in their geographic range and abundance, resulting in changes to community composition, genetic variation and biotic interactions (Davis 1983, Hewitt 2000, Gill et al. 2009). For example, in eastern North America and Europe, glaciers and intense cold forced many species southward during the glacial periods (Hewitt 2000), though the degree of isolation and latitude of refugia may have differed between continents (Lumibao et al. 2017). In contrast, in California and South America, populations contracted into multiple, climatically stable refugia (Grivet et al. 2006, Carnaval et al. 2009). While such broad characterizations of historical demography are informative, the details remain elusive. Open ecological questions include the following: with what speed and trajectory did species change their distributions? To what degree did long distance dispersal contribute to range shifts? What factors limited species ranges? What traits determined the degree of change a species experienced? Did closely-related species, or species in the same communities, react similarly? Open evolutionary questions include: How large and numerous were refugia? To what degree did refugia experience genetic divergence? What changes in adaptive and neutral genetic diversity occurred during recolonization? Answering these questions requires quantifying key facets of biogeographic history including location of refugia, population sizes, migration rates and timing of demographic changes.

Resolving Quaternary biogeographic history is important not only for fundamental ecological and evolutionary science, but is also crucial for addressing applied scientific questions about species' response to contemporary climate change (Barnosky et al. 2017, Nogués-Bravo et al. 2018). Species geographic ranges, demography and community composition are already changing under climate change, which will impact ecosystem services, carbon cycling/carbon storage, and global biodiversity. Humans may be able to mitigate some socioeconomic and ecological impact if we can understand, forecast, plan and react to ongoing shifts. Prediction and mitigation of species' response will depend upon sound knowledge of the speed at which species can disperse into new habitat, the degree to which species will respond in synchrony and maintain community structure, the length of time that species can persist at small population sizes in suboptimal habitat, and the traits that limit or enhance species' successful response (Petit et al. 2008, Woolbright et al. 2014). Such knowledge

Box 1. Approximate Bayesian computation

A major challenge to integrating disparate types of biogeographic data is that each represents the outcome of different ecological and measurement processes. As a result, measurement and process uncertainty differ among data types. Bayesian statistical approaches provide a natural framework for incorporating different sources of information and accounting for multiple sources of uncertainty. Approximate Bayesian computation (ABC) is a method for statistical inference in situations in which it is difficult or impossible to articulate the likelihood of the model (Bertorelle et al. 2010). As an illustration of a simple ABC analysis, consider the two population divergence models presented in Fig. 6. In ABC, a generative model that describes the process of interest (i.e. population divergence, Fig. 6A) is used to simulate datasets whose parameter values are randomly drawn from user-defined prior distributions (Fig. 6E–G). Summary statistics (Fig. 6B–C) are calculated for each simulated dataset (e.g. individual genotypes) and recorded along with the parameter values used to generate the dataset. Examples of these statistics may include F_{ST} in genetic studies, the tree topology in phylogenetics (Jackson et al. 2017), or community composition in ecology. Choice of summary statistics is an important aspect of ABC, as the summary statistics must reflect changes in the underlying parameters driving the simulations; there is not a single set of summary statistics that will be informative for all studies (a point we revisit in section 3, Advance 1). Summary statistics from the simulated datasets are then compared to summary statistics calculated from an observed dataset, and the difference between observed and simulated statistics is quantified using a distance metric. This process of simulation and computation of a distance metric is repeated many times (e.g. $>10^6$). For each repetition, inputs (i.e. parameter values) and outputs (i.e. summary statistics) are recorded. Simulations that produce summary statistics closest to those calculated from the observed data are retained (typically 0.1–3% of all simulations, with authors often testing multiple thresholds to ensure results are not sensitive to the precise choice, Fig. 6D; see also Bertorelle et al. 2010) based on the assumption that they best reflect the true process. These retained parameter values are collectively used to form posterior distributions for relevant parameters (e.g. dispersal components, population size changes, refugium locations; Tavaré et al. 1997, Pritchard et al. 1999, Beaumont 2010, Fig. 6E–G). In the example in Fig. 6, two summary statistics provide a strong signal for estimation of two model parameters (Fig. 6E–F) and a poor signal for a third (Fig. 6G).

Since its introduction (Tavaré et al. 1997, Pritchard et al. 1999), ABC has been widely used in population genetics to study biogeographic processes including invasion (Estoup et al. 2001, 2010, Benazzo et al. 2015), historical demography (Thornton and Andolfatto 2006, François et al. 2008, Rougemont and Bernatchez 2018) and the role of environmental change and geological processes in determining biodiversity (Patiño et al. 2015, Barnes and Clark 2017). The ABC statistical framework is flexible; it can be applied to any system that can be described using a generative model and for which a set of informative summary statistics can be identified. ABC studies in ecology use a variety of modeling approaches for the focal system such as a system of difference or differential equations, agent-based models, or cellular automata. In population genetic studies, a reverse-time coalescent simulation model is parameterized by an explicitly-defined set of demographic events (e.g. population divergence or size change), and outputs a set of individual genotypes. However, it is possible to couple coalescent simulations with detailed agent-based forward-time simulations of population dynamics (in which only demographic parameters are set by the user). With this formulation, demographic history (i.e. all colonization events and population sizes through time) is an output of the stochastic agent-based model (rather than defined a priori). This demographic history then informs the more computationally efficient coalescent simulation. Agent-based simulations are computationally expensive but facilitate more realistic spatial and demographic processes (Hoban et al. 2012). At the same time, algorithms are constantly becoming more efficient and ABC simulation lends itself to easy parallelization on ever-growing compute clusters, both of which offset the computational costs of model realism. Nonetheless, there are limits to complexity of models and scale of resolution (as with any inference methods).

One main advantage of ABC is the ability to make inference about complex processes such as biogeographic histories (species' populations moving across space) which often include multiple bottlenecks, spatial expansions, changes in migration rate, etc. Another advantage of ABC is that quantitative evaluation of confidence in model selection and parameter estimation is built into the framework via the analysis of pseudo observed datasets, which provides an assessment of the confidence in model selection and the accuracy and precision of parameter estimates (Bertorelle et al. 2010). In addition, the underlying demographic and genetic models for ABC are built on well-understood processes (e.g. migration, coalescence). Complex models can be constructed by adding together many simple components at varying spatial and temporal scales (Bertorelle et al. 2010, Csilléry et al. 2010, Hoban 2014). A final benefit of ABC (elaborated on in section 3) is the ability to incorporate prior information from different sources – a natural way of combining existing information with new data. However, in practice, ABC analyses in population genetics often rely on broad, uninformative prior distributions (Fig. 6E–G), or priors defined by intuition and the investigator's interpretation of available knowledge, missing an opportunity to leverage data from other fields in a quantitative fashion. Future developments in ABC will likely benefit from combining non-genetic data with genetic data such that both are used in the simulation and analysis.

can also assist in designing protected area reserve networks and choosing species for restoration (e.g. species with robust climate tolerances) or translocation (e.g. those species that are unable to move on their own). In addition, knowing the distribution of rates of species' movement provides a sense of overall risk for all species. Of course, contemporary habitat fragmentation will present species with new challenges not

faced in the past (i.e. the Holocene); nonetheless, knowledge about how species and communities respond to changing conditions is foundational for ecosystem management in the future. Finally, a better understanding of geographic and genetic structure during glacial periods can help identify regions and environments that have supported species' survival (which may be informative for placing contemporary

protected areas), and regions which could harbor high genetic variation today. These regions may also be useful reservoirs of genetic resources for seed banks, breeding or translocation (Fady et al. 2016).

To understand species' biogeographic histories, different scientific communities have used various data types and methods for translating data to inferences. In this review we focus on three approaches: 1) ecological niche models (ENMs), in which species' occurrences (contemporary and/or historical) are combined with environmental data to estimate ecological niches; 2) paleoecological approaches, in which fossils (especially fossil pollen, wood fragments, bones or shells) are used either directly or in a pollen-vegetation model to track species ranges through space and time; and 3) genetic/phylogeographic approaches, in which ancient or modern DNA sequence data are combined with coalescent or phylogeographic models to quantify demographic change. In discussing these approaches, we will focus particularly on plants, while including animal examples when possible. The reasons for a plant focus include: 1) there are extensive networks of well-curated fossil pollen records; 2) new advances in pollen-vegetation modeling now make it possible to translate these pollen records to vegetation abundance; and 3) chloroplast DNA from plants is inherited via the unit of colonization (seeds), and thus provides a long-lasting record of colonization history. Nevertheless, the concepts and methods we discuss also apply to historical data types indicative of presence of any taxa (e.g. packrat middens, shell collections, macrofossils, historical human accounts, archaeological materials, etc.).

In each of these three approaches, early inferences were often based on the visual interpretation of mapped data (e.g. DNA haplotype distributions or clines of percentage of fossil pollen). In subsequent decades, major breakthroughs in the size and resolution of datasets, aggregated databases, and new statistical analyses have transformed each of these disciplines. While the use of multiple datasets is on the rise (Gavin et al. 2014), many investigations still use only, or primarily, one data source (e.g. only genetics) and methods for true integration across approaches are lacking. At best, results of different strands of analysis are qualitatively compared (see examples in Gavin et al. 2014). Fundamentally, each data type reflects different aspects of species' biogeographic history, and thus intuitively no single data type alone is sufficient to reveal all aspects of a species' biogeographic history. This is because each type of data is the outcome of different underlying processes (e.g. mutation, dispersal, demography, sedimentation, etc.), reflects different spatial and temporal scales, and has recognized limitations and errors. Each data type also typically has different spatial coverage of a species' range, and different abundance and resolution (e.g. all data types have collector bias; fossil pollen sites and specimen records are more abundant than genetic samples). As a consequence, importantly, inferences from analyzing different data types often conflict with each other. Of course, all statistical estimation will have associated error, and the objective is to minimize the error.

However, even after decades of work, achieving sufficient resolution of species' biogeographic histories (i.e. where species were, when and why) remains a challenge (Feurdean et al. 2013, de Lafontaine et al. 2014, Fuentes-Utrilla et al. 2014).

We assert that an analytical revolution is needed to resolve ongoing disagreements in biogeographic understanding. Specifically, a true advance will require quantitative integration of the data and models used in different disciplines. The biogeographic community has recognized the need for a comprehensive framework to combine information from multiple approaches and data types (Fordham et al. 2013, Gavin et al. 2014, Collevatti et al. 2015a), and some progress has been made. Cutting edge approaches (section 3) include the integration of dispersal and demographic processes into niche models (Pagel and Schurr 2012) or the use of one data type to define conditions or priors for a model informed by another data type (Espíndola et al. 2012, Alvarado-Serrano and Knowles 2014, Brown et al. 2016). Nonetheless, an approach that truly integrates all data sources has yet to be developed.

An integrative framework should 1) leverage data and understanding from different fields; 2) model the key relevant processes that produce the data (not only demographic processes but also mutation, pollen deposition, etc.); 3) quantify uncertainty within and among data types and processes; and 4) be computationally tractable and well-validated (Schaub and Abadi 2011, Fordham et al. 2013, Collevatti et al. 2015a). The development of new, integrated informatics approaches will allow for improved understanding and quantification of fundamental aspects of biogeographic history mentioned above (i.e. speed and timing of migration, location and size of refugia, degree of long distance dispersal, etc.). Ideally, these integrative approaches would account for the various sources of uncertainty inherent in the data and processes. This characterization of uncertainty will allow for improved accuracy in parameter estimates and an understanding of statistical confidence. Additionally, bias will be reduced compared to analysis of single data types (Schaub and Abadi 2011, Talluto et al. 2016). This will ultimately help achieve biological understanding by resolving the above-mentioned fundamental, long-standing questions regarding how species and communities responded to Quaternary climate change. Integrative advancements will be enabled by recent improvements in data and models, including the following: the availability and curation of large biological and environmental datasets (e.g. Sequence Read Archive, Neotoma Paleoecology Database, Global Biodiversity Information Facility, GeOME, CHELSA, WORLDCLIM, etc.), new pollen-vegetation models for reconstructing past species' ranges and forest communities (Dawson et al. 2016), advances in spatially-explicit demographic modeling (Ray et al. 2010), availability of high quality ancient DNA (Llamas et al. 2016), identification of new and informative spatial-genetic summary statistics (Peter and Slatkin 2013, Alvarado-Serrano and Hickerson 2016, 2018), and joint demographic-niche models (Pagel and Schurr 2012).

In this article, we will: 1) describe the major data sources, including their strengths and weaknesses, as well as approaches used to infer biogeographic history from each in isolation (section 1); 2) review the current state-of-the-art of integrative approaches, focusing in particular on approximate Bayesian computation (ABC) (section 2); 3) propose a novel set of scaffolded and feasible modeling advances to aid in integrating these data more tightly and in a more quantitative manner (section 3); and 4) provide an outlook on future opportunities and computational and technical feasibility (section 4). Readers interested in this topic may also turn to recent, complementary papers. Fordham et al. (2013) focus on practical applications of forecasting models that use multiple data sources for biodiversity conservation, e.g. predicting biodiversity loss or determining species' endangerment status; Maguire et al. (2015) explain the complementary use of fossil and environmental data for delineating a species' niche and testing for niche change over time; Alvarado-Serrano and Knowles (2014) focus in particular on integrating genetic and ecological niche models; and finally, Gavin et al. (2014) focus on post hoc methods for combining data sources (e.g. analyzing data separately and comparing conclusions by overlaying maps).

1. Introduction to data types and approaches

Different kinds of evidence contain different information into species' biogeographic history including: georeferenced modern and historic natural history collections; geological and climatic evidence of changes (e.g. glacial advances and retreats); pollen, charcoal and macrofossils; archeological evidence (e.g. evidence of human use of particular species); modern measures of traits, dispersal and demographic processes; and modern and ancient DNA data (Fig. 1). These different sources of evidence (Fig. 1) are influenced by a combination of shared and distinct processes, and consequently each can provide complementary and unique knowledge about biogeographic history. For example, georeferenced collections data can be used to calibrate niche models for hindcasting habitat suitability which provides information about environments where species could potentially survive. Fossil evidence, interpreted through appropriate models and temporal calibration, provides physical documentation of species presence and relative abundance. Genetic data can be used to estimate temporal changes in abundance and migration between populations. Because each is influenced by different, but overlapping, processes (Fig. 2), these data are heterogeneous and incomplete and vary in spatial and

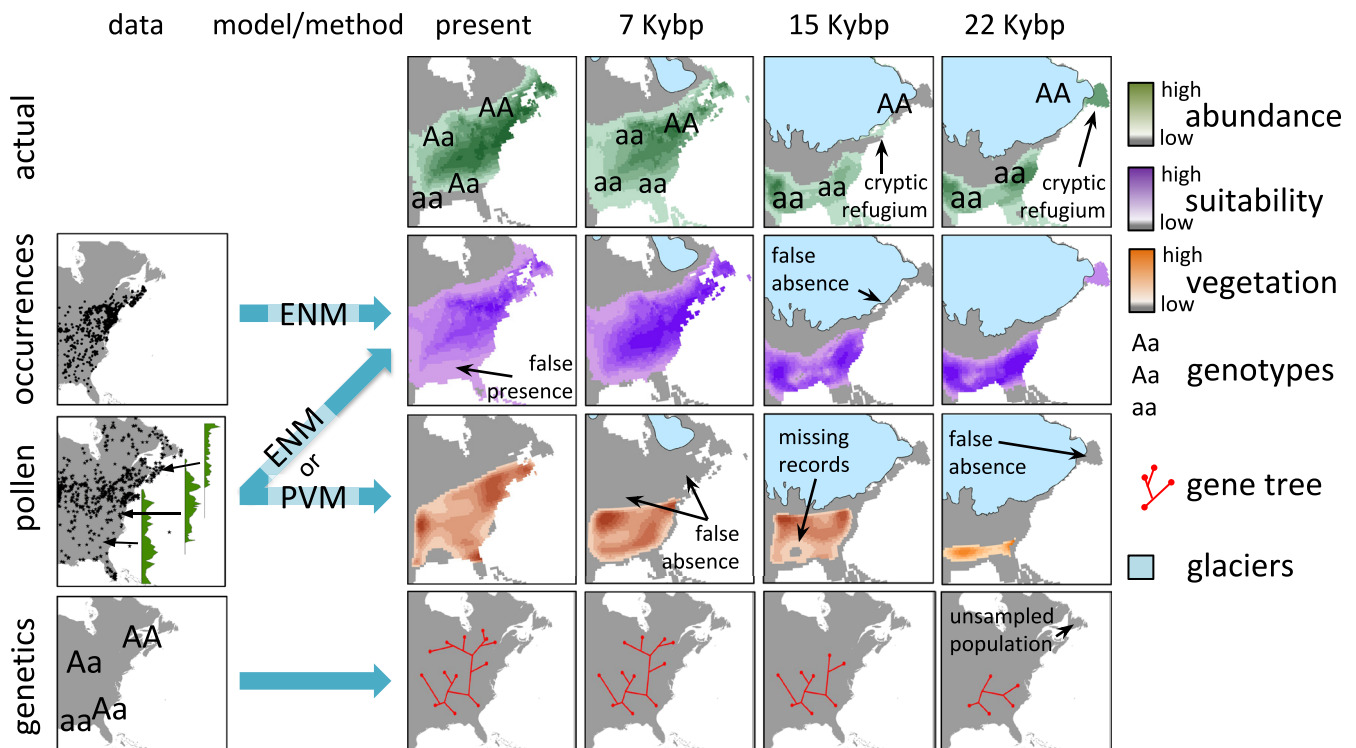


Figure 1. Conceptual summary of the data types and methods used to infer species' biogeographic history. The top row shows a hypothetical (unobserved) distribution, abundance and genetic trajectory of the species through time. The leftmost column illustrates different types of data (observations) used by the respective methods (ecological niche modeling combined with global circulation models, pollen-vegetation modeling, genetic analysis) to infer biogeographic history (right four columns). Note that fossil pollen can be modeled using either an ENM or a PVM. Each data type and method reflects part of the real history, but also suffers from data gaps, issues with false positives and negatives, and limited ability to depict all relevant aspects of biogeographic history (note the black arrows). The examples are for a simulated species whose characteristics and history are based on green ash *Fraxinus pennsylvanica*.

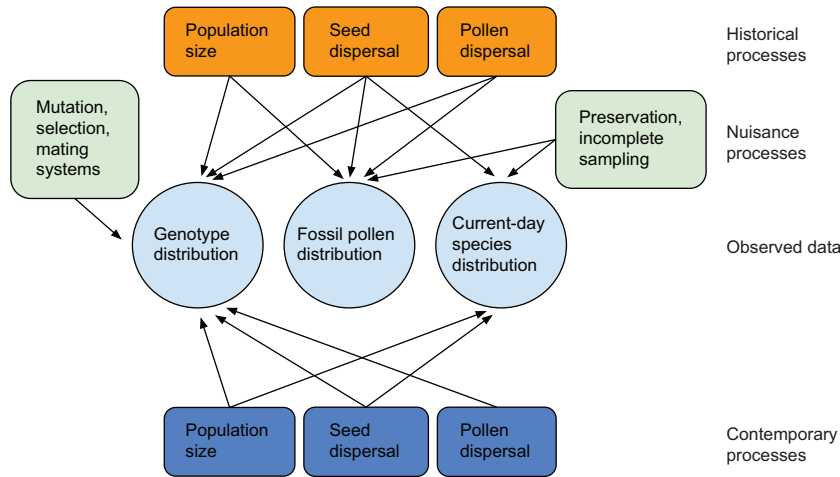


Figure 2. The directional influences of major processes on the observations of three data types used for historical biogeographic inference. Circles represent observed evidence, rectangles represent generative processes, and arrows represent causal relationships. Orange and blue boxes correspond to historical and contemporary processes whereas green rectangles describe processes that influence observed data, but that are not of immediate concern when estimating range-shifts. While ‘nuisance processes’ are enormously important in other realms of ecology and evolutionary biology, here they represent processes that typically do not directly influence biogeography. Note that while genetic and occurrence data continue to be influenced through contemporary processes, fossil evidence reflects only historical processes.

temporal resolution (Table 1). The raw data themselves must be collated, cleaned, and often used in concert with modeling approaches and additional datasets (e.g. environmental data, Maguire et al. 2015). Biologists have worked for decades with these approaches to infer species’ past ranges, resulting in key discoveries (e.g. pollen diagrams have shown that *Pinus* was among the first postglacial colonizers in North America (Davis 1983), while Petit et al. (2003) used population genetic data to document isolated glacial refugia in southern peninsulas and mid-latitude ‘mixing’ of lineages in Europe). Nonetheless, recent work demonstrates that current understanding of biogeographic history is lacking, and sometimes misleading, due to increasingly-recognized issues with data and models in respective disciplines. For example, rates of postglacial forest movement inferred by molecular marker

data are two to five times slower than rates inferred from pollen data (McLachlan et al. 2005, Cheddadi et al. 2014), and both are faster than expected based on observed seed dispersal (Reid 1899, Clark et al. 1998). This conflict, and broadly the inability of any one data type to unambiguously resolve biogeographic history, means that the use of only one approach limits our collective understanding of evolutionary and community dynamics.

Inference of biogeographic history is an inherently difficult problem. The processes of interest occur at scales that vary by several orders of magnitude (e.g. from meters to hundreds of kilometers, and hourly to millennial), and data coverage is inconsistent and uneven across space and time (e.g. genetic data collection is typically not systematic, pollen records from unglaciated regions during the last glacial

Table 1. Summary of major data types and their predominant advantages and disadvantages (IE – indirect evidence; DE – direct evidence). See Glossary for definition of other terms and abbreviations.

Type	Process	Common method of analysis	Advantage	Disadvantage for biogeographic inference
Fossil pollen occurrence	Presence of a species in a region (DE)	PVM, sometimes ENM	Moderate spatial and temporal accuracy	Pollen can disperse far from a population; abundance is relative to other species
Macrofossils occurrence	Presence of a species in a locale (DE)	ENM, sometimes ABC	Good spatial and moderate temporal accuracy	Are relatively rare
Modern DNA	Demographic and spatial change (IE)	ABC simulations	Reveals non-relative abundance especially low abundance	Decreasing resolution through time; some processes produce same pattern
Ancient DNA	As modern DNA and fossils (IE/DE)	In development	As above but more spatial and temporal accuracy	Technically difficult; rare and thus coarse resolution; error prone
Contemporary observations, historical records	Suitable habitat (IE) at a point in space and time	ENM, sometimes ABC	Bounds where a species can likely exist; precise	Many assumptions (static niche, etc.); cannot distinguish sinks/sources

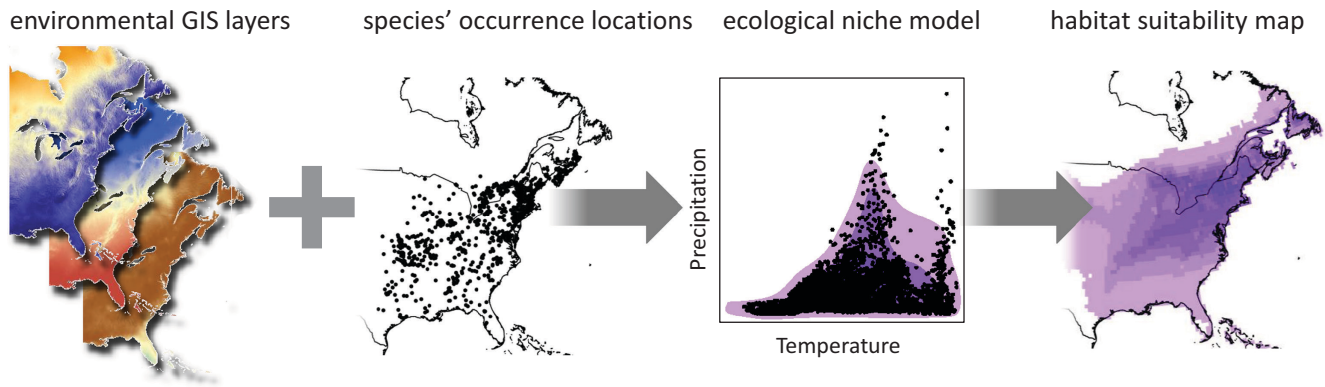


Figure 3. Environmental and specimen occurrence record data can be used to construct an ecological niche model (ENM). The input data consists of matching environmental data (left) representing climatic conditions, soil and geology, and other factors with occurrence records (second from left). The model then identifies conditions most conducive to presence of the species (second from right). This model can then be projected back onto a landscape using contemporary environmental conditions (right) or to environmental layers for different time periods.

maximum are sparse due to the lack of glacial lakes, occurrence records may show collector bias, etc.). In addition, there are numerous parameters of biogeographic history that need to be simultaneously inferred, i.e. species' abundance at key time points; process parameters such as proportion of long distance dispersal events; location(s) of glacial refugia; and the timing and rate of post-glacial expansion. Due to these considerations as well as data-specific attributes, benefits and limitations, we expect that joint biogeographic inference based on all data types will reduce gaps in information and knowledge and reduce uncertainty and bias resulting from inference based on a single data type (Schaub and Abadi 2011). Before discussing recent and proposed developments towards integration of information, we begin by summarizing the following: 1) environmental data and specimen occurrence records; 2) fossil (especially pollen) data; and 3) genetic data. For each of these data types, we describe how it is obtained, measured and analyzed; its scale; corresponding key assumptions, weaknesses and strengths; and some successful applications.

Environmental and specimen occurrence data

Contemporary occurrence data (geolocation for each individual, ideally with a stated accuracy) can be obtained from modern global repositories like the Global Biodiversity Information Facility (GBIF 2018; though extensive cleaning and checking may be required), while environmental data for those same geolocations (e.g. minimum temperature, summer rainfall, etc.) may be obtained from a growing number of environmental databases like CHELSA (Karger et al. 2017) and WORLDCLIM (Fick and Hijmans 2017). Ecological niche models (ENMs, i.e. species distribution models) are used to relate information on occurrence and environmental data. Depending on whether presence/absence or presence-only data are used to train the model (Guillera-Aroita et al. 2015), they estimate either the probability of occurrence or habitat suitability by identifying correlations between species'

geo-located occurrence records and environmental data at those locations (Fig. 3). ENMs can be trained with contemporary data or macrofossils and datasets from paleoclimate simulations (Nogués-Bravo 2009), but as we explain below, pollen fossil occurrences can be interpreted through a separate modeling framework which takes into account the processes producing the data. The spatial grain of the model depends on the spatial resolution of the environmental data; the finest grain obtainable for contemporary continental-scale climate data is usually ~1 km (Fick and Hijmans 2017, Karger et al. 2017), although spatial uncertainty in occurrence data coordinates may require using coarser environmental data. Climate reconstructions from paleoclimate models can be downscaled to the same resolution as contemporary environmental datasets, but typically the 'native' resolution at which it is modeled is 50–200 km. ENMs are frequently used to 'hindcast' a species' likely historical geographic distribution by projecting the model to environmental layers for historical periods (Guralnick 2007, Nogués-Bravo 2009, Roberts and Hamann 2012). This approach has helped to answer diverse questions about biogeographic history, e.g. locating refugia to which species retreated during past periods of environmental stress (Carnaval and Moritz 2008, Roberts and Hamann 2015) and paths of subsequent expansion (Forester et al. 2013); understanding the relative importance of drivers of historic range shifts (Smith 2013, Santos et al. 2017); and identifying evolutionarily significant yet cryptic taxonomic units (Phuong et al. 2014).

Despite their utility, use of ENMs for historical inference is predicated on the key assumptions that 1) the species' spatial distribution is in equilibrium with its environment during the period in which the model is trained and projected (Wiens et al. 2009, Peterson et al. 2011, Maguire et al. 2015); 2) the fundamental niche does not evolve across time; 3) there is no local adaptation to environmental conditions (Valladares et al. 2014, Maguire et al. 2015, Smith et al. 2019); and 4) the environmental data to which species' occurrences

are matched, and to which the model is projected, reflects the key factors driving distribution (Franklin et al. 2013) and has relatively low error (Varela et al. 2015). As a result, correlative ENMs do not account for phenomena such as ecological sinks, dispersal limitation, low population sizes, or missing mutualists such as pollinators (Franklin 2010, Guisan and Thuiller 2005). ENMs are also limited in their ability to predict across time periods experiencing dramatic environmental change because current distributions (and thus the model) may not encompass all environments inhabitable by a species (Jackson and Overpeck 2000, Barve et al. 2011, Maguire et al. 2016). A further challenge is spatial estimation of microclimates in areas with variable topography such as mountains. Lastly, projections to the past using ENMs are challenged by variability among different historic global climate models (Varela et al. 2015). Current developments in niche modeling seek to address some of these issues by incorporating dispersal (Midgley et al. 2010, Pagel and Schurr 2012, Thuiller et al. 2013) and biotic interactions (Ovaskainen et al. 2010, Anderson 2017), correcting for sampling bias and error (Velasquez-Tibata et al. 2016, Hefley et al. 2017), and accounting for local adaptation (Smith et al. 2019).

Paleoecological data

Fossil pollen and macrofossils are biological artifacts deposited and preserved in sediment. Samples that are taken from a sequence of sediment depths (with depth as a proxy for age as sediment is laid down over time) provide information, with varying degrees of certainty, about past vegetation communities. Each sediment record represents changes in vegetation at different spatial scales. The spatial resolution of the vegetation signal in a sediment record depends on many factors, including the size of the sedimentary basin, as well as landscape structure and composition. Despite this variable spatial resolution, when networks of paleoecological records are considered collectively, it is possible to make reasonable inference on shifts in species' ranges and relative abundance over space and time. These large-scale analyses are facilitated by recent advances in cyberinfrastructure and data reproducibility, especially the digitization and curation of thousands of sediment cores in high-quality and easily accessible databases (e.g. Neotoma Paleoecology Database; Williams et al. 2018). To associate calendar dates with paleoecological samples, age–depth relationships are estimated based on stratigraphic constraints that typically include radiometric or biostratigraphic dates. As a result of this process, sample ages are not known with certainty, though recent advances in Bayesian age–depth model development allow for more robust estimation of sample age (Blaauw and Christen 2011, Parnell et al. 2011).

Paleoecological data is arguably the best evidence of past vegetation change since it provides physical confirmation of species existence and relative abundance. Pollen data itself has revealed phenomena including the mid-Holocene decline of eastern hemlock (*Tsuga canadensis*; Zhao et al. 2010)

and shifts in the northeastern US prairie–forest boundary (Williams et al. 2009). For more than a century, fossil pollen and macrofossils (e.g. seeds or wood) have been used to infer past distributions and abundances of vegetation. Early biogeographic inference from paleoecological data was based on the qualitative interpretation of raw pollen percentage maps. More recently, especially in the last several decades, quantitative methods have been applied in paleoecology to make inference about past vegetation distributions. First, quantitative approaches were developed to account for the non-linearity of the pollen–vegetation relationship; these methods included the R-value method (Davis 1963, Parsons and Prentice 1981) and the modern analogue technique (Overpeck et al. 1985). While the modern analogue technique has been widely used, it is described as a more qualitative approach, and does not account for biological and physical processes that govern pollen–vegetation relationships. More recent quantitative approaches build upon the process-based ideas behind the R-value model, and formalize theoretical understanding of the key processes that link vegetation and sediment pollen. These key processes minimally include pollen productivity, dispersal and sedimentation. These approaches, referred to as pollen–vegetation models (PVMs; Fig. 4), include LOVE/REVEALS (Sugita 2007a, b) and the more recently developed STEPPS (Dawson et al. 2016). Both PVMs formalize understanding of key processes, though LOVE/REVEALS is more widely used. However, LOVE/REVEALS requires input parameters describing atmospheric conditions and pollen characteristics, and does not account for spatial or temporal dependence. As a result, with the LOVE/REVEALS model, vegetation can only be estimated for locations and times for which there is pollen data. The more recently developed PVM STEPPS is a Bayesian hierarchical model that improves upon existing methods by 1) accounting for spatio-temporal dependence, 2) not requiring parameter inputs, and 3) allowing for the estimation of past vegetation with explicit accounting of uncertainty. STEPPS was recently used to reconstruct forest composition for the upper midwestern United States for the last two millennia (Dawson et al. 2016). We note that fossils, including fossil pollen, can be used as input to ENMs, but here we discuss them primarily via the PVM because (a) the PVM includes the processes influencing pollen fossil occurrence while using pollen counts in an ENM would not account for this, and (b) use of the fossil pollen in multiple places in the framework we propose below would equate to using this data source twice, perhaps over-inflating its influence on inferences.

Despite the advances in PVMs and cyberinfrastructure, paleoecological data presents several challenges. These include poor taxonomic resolution (i.e. inability to identify pollen to the species level, rather than only to genus level in species-rich genera like *Quercus*, Davis 1983), temporally varying pollen–vegetation relationships (Kujawa et al. 2016), and both false-positive and false-negative errors (i.e. evidence of species' presence in a location where it was not; lack of

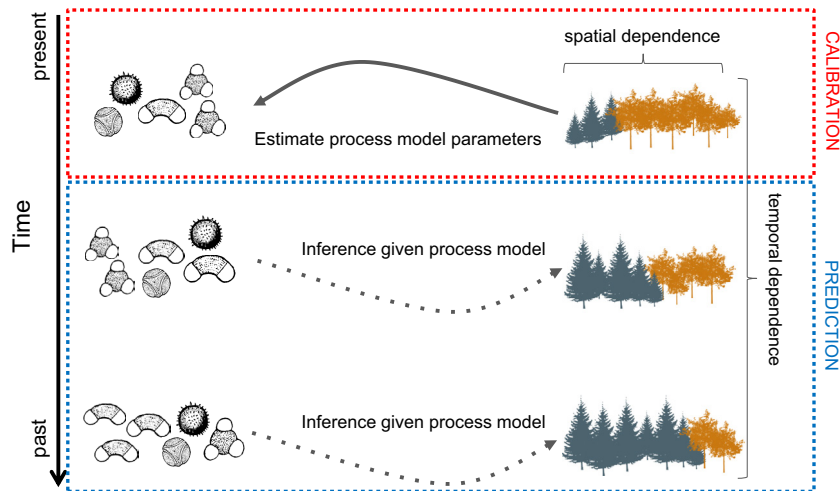


Figure 4. Conceptual schematic of a pollen–vegetation model (PVM). Two stages are identified: calibration and prediction. In the calibration stage, the parameters that describe the key process that link pollen to vegetation are estimated. In the prediction stage, parameters estimated during calibration are used to predict vegetation from pollen data back through time. As noted in the figure, spatial and/or temporal dependence exist in the data; some PVMs account for this dependence while others do not.

evidence in a location where the species truly was, sometimes called ‘cryptic’ refugia). For example, present day locations of low-density American beech *Fagus grandifolia* populations are not evident based only on pollen occurrence in the sediment, suggesting that low-density historic populations may also fail to be identified (McLachlan and Clark 2004). This problem is further complicated by the appearance of beech pollen found in sediment hundreds of kilometers from any known populations (McLachlan and Clark 2004). Lastly, pollen shows latitudinal bias in preservation as a function of the distribution of glacial lakes; in North America this translates to more northern records. As a result, many sites at lower latitudes that may have served as refugia or expansion corridors are poorly represented in these data, which is a major challenge for studies seeking to identify these locations.

Genetic data

Genetic data are obtained from living (or once living) individuals. For analysis, these data are often aggregated to frequencies of alleles at a population level. These population genetic data have been used in traditional phylogeography to visualize distinct genetic clusters, or haplotypes, on a map (Avise et al. 1987). Divergent lineages can provide evidence of distinct, isolated refugia, while high diversity may lead to conclusions about centers of endemism, refugia and other phenomena. For example, chloroplast DNA haplotypes revealed southern Ice Age refugia in European oaks (Ferris et al. 1993) and elevational, rather than latitudinal, migration in California oaks (Grivet et al. 2006). As with pollen data, early interpretation was entirely visual, but the past two decades have seen development of model-based methods (e.g. the software IMA, Nielsen and Wakeley 2001; the software BEAST, Drummond and Rambaut 2007) for inferring migration rates or demographic dynamics, from genetic data. For example, Bayesian skyline plots (a method to infer

effective population size through time) revealed multiple expansions and contractions in musk ox over 60 000 years (Campos et al. 2010). More recently, the desire for more flexible models and better accounting for uncertainty has led to the use of approximate Bayesian computation (ABC, see Box 1) to estimate migration parameters and track colonization and invasion routes (Estoup et al. 2010).

Genetic data retain signatures of past demographic and selective processes, including strong signals of changes in population size and migration (Milligan et al. 1994, Csilléry et al. 2010). A century of development of population genetics theory has improved inferential power, e.g. cane toad *Bufo marinus* invasion rates as inferred from genetic estimates were 50 km year⁻¹, matching estimates from field observations of 55 km year⁻¹ (Estoup et al. 2010). Genetic data are often easy to collect in contemporary organisms (e.g. non-invasive sampling of hair, feathers, or feces) and sometimes available for calibrated dates in the past (e.g. ancient DNA). Additionally, the quantity of DNA that can be sequenced in a typical study has expanded by two to five orders of magnitude in recent years, increasing the precision of parameter estimates. Typically, modern DNA is analyzed, but in the past decade analysis of ancient DNA (sometimes thousands of years old) has become feasible (de Bruyn et al. 2011).

A major disadvantage of biogeographic inference based on genetic data is that most genetic models do not include aspects of realistic biogeography that may be key for making inferences. They often ignore variability through time, space and among individuals in demographic and migration processes (e.g. models often assume constant migration when migration events may occur in distinct bursts, Marino et al. 2013). While rapid developments in sequencing technologies have allowed researchers to survey far larger proportions of the genome, using these data to estimate parameters in

complex demographic models remains a challenge. Inferences drawn from model-based phylogeography are inherently limited by the number of models considered in the analysis and the degree to which they reflect the real world. New methods to best utilize computational resources and genomic data are still being developed, but have somewhat lagged behind the explosion of genomic data (Robinson et al. 2014, Shafer et al. 2015). A continual challenge in genetics-based inference is that very different demographic processes can sometimes lead to similar genetic signatures (e.g. fragmentation and population bottleneck), leading to ‘false signals’ and identifiability issues (Petit et al. 1997, Gong et al. 2008). For example, range expansions create signals easily confounded with adaptation (Beaumont and Nichols 1996, Lotterhos and Whitlock 2014), while long distance colonization from one refuge has a pattern similar to multiple refugia (Bialozyt et al. 2006). Lastly, genetic data are (usually) limited to extant lineages – thus loss of alleles results in lost information. Ancient DNA holds great promise to alleviate this issue. However, ancient DNA is quite rare, is only preserved in certain suitable locations, and when present is typically restricted to organelle DNA (chloroplast or mitochondria, Parducci et al. 2012). Compared to nuclear DNA, organelle DNA may show strong bias (e.g. from sex-biased dispersal) and only represents one genealogy and thus has low information content.

2. Recent advances towards integration

As described above, approaches for inference of past range shifts based on each data type individually have seen substantial methodological developments and expanded application in both paleobiology and paleoecology. Known weaknesses and conflicting conclusions from different approaches indicate that combining information will be necessary for true advancement. However, few studies simultaneously integrate information from genetic, paleoecological and environmental/occurrence data (Svenning et al. 2011). While an increasing number of studies do consider two lines of evidence (e.g. ENM and genetic), such studies only assess agreement qualitatively (reviewed by Alvarado-Serrano and Knowles 2014, Collevatti et al. 2015a). Typical practice is to analyze different data types separately (e.g. a genetic dataset and a niche model), then base ultimate inference on visual, non-quantitative comparison of results (Fordham et al. 2013, Gavin et al. 2014, Bagley et al. 2018). Another approach involves using ENMs to develop hypotheses to be tested using phylogeographic data (Richards et al. 2007, Forester et al. 2013, Alvarado-Serrano and Knowles 2014, Collevatti et al. 2015b, Roberts and Hamann 2015).

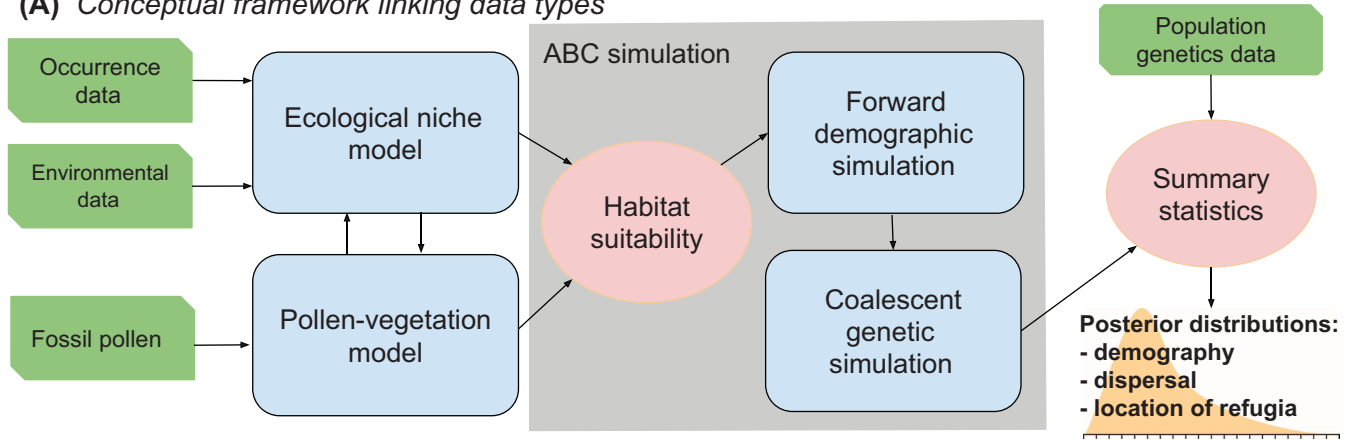
Recently, investigators proposed using ENM output (i.e. predictions of the niche) to inform simulations for ABC-based biogeographic inference (He et al. 2013, Brown et al. 2016, He et al. 2017). Specifically, in these simulations, ENM output was used to define habitat quality. Carrying capacity was thus allowed to vary as a function of changing habitat quality (Brown and Knowles 2012). Thus, an ENM-based habitat

suitability surface ties demography directly to habitat, at a much finer temporal and spatial resolution than previously achieved. For example, Brown et al. (2016) used ENMs to create habitat suitability maps for time slices throughout the last 50 000 years for the perennial herb *Penstemon deustus*. The challenge was to then translate these habitat suitability maps to demographic parameters (i.e. population carrying capacity, habitat ‘friction’). To determine the appropriate translation, a set of possible transformations (linear, sigmoid, etc.) were tested (Brown et al. 2016), and the translation that best fit contemporary data was selected. The authors employed ABC-based analyses (Box 1) in which demography was constrained by the habitat matrix, using the best fit translation, to infer parameters governing ancestral population size, maximum carrying capacity, and migration rates. They then used these parameters to forecast future population trajectories for the next century. Others have recently used ENM-informed ABC to determine relative support among models of post-glacial population fragmentation (Dellicour et al. 2014, Massatti and Knowles 2016). As ENMs are an accessible and popular method, we expect more investigators to follow these examples in the future.

Similar proposed approaches (Richards et al. 2007, Knowles and Alvarado-Serrano 2010, Fordham et al. 2013) have only been applied in a handful of empirical studies (Knowles and Alvarado-Serrano 2010, He et al. 2013, 2017, Brown et al. 2016, Massatti and Knowles 2016) partly due to a lack of a comprehensive informatics framework and infrastructure. Moreover, ongoing work is making strides but 1) takes ENM-derived habitat suitability as ‘known’ then simulates population and genetic processes without accounting for uncertainty in the ENM fit (Brown et al. 2016), 2) has prohibitive data requirements (dispersal and multi-site demographic data from many generations at each site, Pagel and Schurr 2012) and/or 3) does not use all three data types. These problems will be explored below. Despite their strengths, state-of-the-art studies lack true quantitative integration of all data types and/or do not handle uncertainty and feedback among steps.

We assert that an ideal framework for integrative modeling for species’ historical biogeographic inference would account for limitations of all available data types/analyses and leverage their strengths. At a minimum, this novel integrative modeling framework would need to: 1) cohesively and quantitatively combine data and understanding from different fields, 2) identify and model key processes, 3) quantify uncertainty in data and processes, and 4) be computationally tractable. Such a framework would have the advantages of using the full range of data available, allowing for the filling in of temporal or spatial gaps, reducing the impacts of data limitations and biases, and quantifying the confidence in conclusions (Schaub and Abadi 2011, Pagel and Schurr 2012, Talluto et al. 2016). In the following sections we describe existing and proposed methods to facilitate full integration of ENMs, fossil pollen and genetic data (ancient or contemporary) in a spatially-explicit analytical

(A) Conceptual framework linking data types



(B) Pictorial representations of data/models

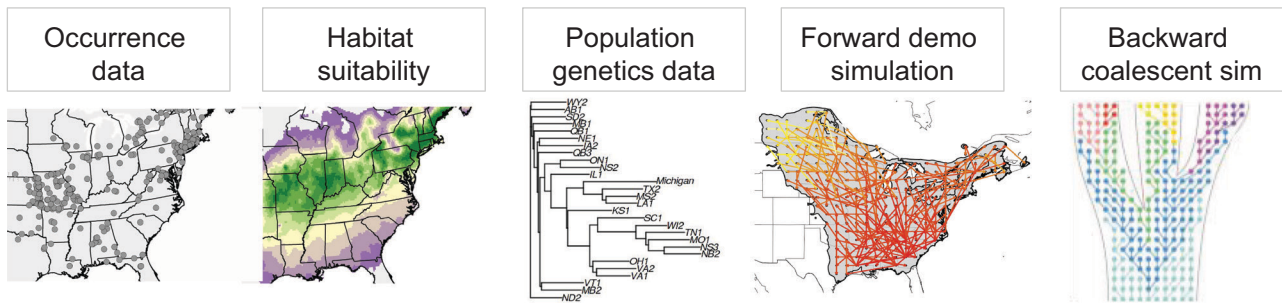


Figure 5. A framework for integrating multiple data types and models into a coherent statistical analysis via the ABC approach to statistical inference of parameters (orange distribution) as described in section 2: Recent advances plus components from Advance 2. Note that we use ‘Environmental data’ to refer simultaneously to contemporary climate data and outputs of paleo global circulation models. Also, the arrows between ‘Ecological niche model’ and ‘Pollen–vegetation model’ signify possible degrees of integration and/or weighting between these boxes; if there was full integration these boxes would merge. (B) Pictorial representation of models/data in (A).

framework (Fig. 5). Several of these advances include integration via ABC (Box 1).

3. Key advances needed for integrative modeling

In this section we describe key advances that are necessary for an integrative approach to biogeographic inference. These advances are presented in order of increasing complexity.

Advance 1. Significant improvements in the state-of-the-art

We begin by suggesting three improvements to better advance the state-of-the-art outlined above: simulation detail, more fully incorporating uncertainty, and identifying/developing informative new summary statistics.

The first improvement is to increase the modeling capabilities of the types of spatially-explicit simulation software used in current state-of-the-art integrative modeling studies (see previous section). The current commonly used software (SPLATCHE, Ray et al. 2010, PHYLOGEOSIM, Dellicour et al. 2014) couple forward-in-time demographic simulations with backward-in-time coalescent simulations to generate genetic datasets across a spatial domain. These software programs allow population presence, size or

connectivity to depend on habitat suitability, which is often determined using ENMs (He et al. 2013, Alvarado-Serrano and Knowles 2014, Prates et al. 2016). This approach has facilitated research to better understand range shifts and fragmentation (Arenas et al. 2012, White et al. 2013, Antoniazza et al. 2014, Dellicour et al. 2014, Mona et al. 2014). However, existing simulation frameworks do not have the ability to sufficiently reflect the fundamental processes in nature that drive biogeographic processes and patterns, especially dispersal syndrome, reproductive strategy, fecundity and life span (Sukumaran and Knowles 2018). Improvements to existing simulation approaches include the specification of complex dispersal kernels and age- or stage-based demographic transition rates (Hoban 2014). Other logical improvements relate to a more realistic characterization of organism dispersal. For plants, this improvement would be to model seed and pollen dispersal separately; for animals, this would be to allow dispersal to vary as a function of life stage (i.e. adult versus juvenile) or sex (given that males and females often have different dispersal propensities). This advance would help ensure that model outputs can accurately represent spatial genetic structure of nuclear and organelle genomes (Petit et al. 2005).

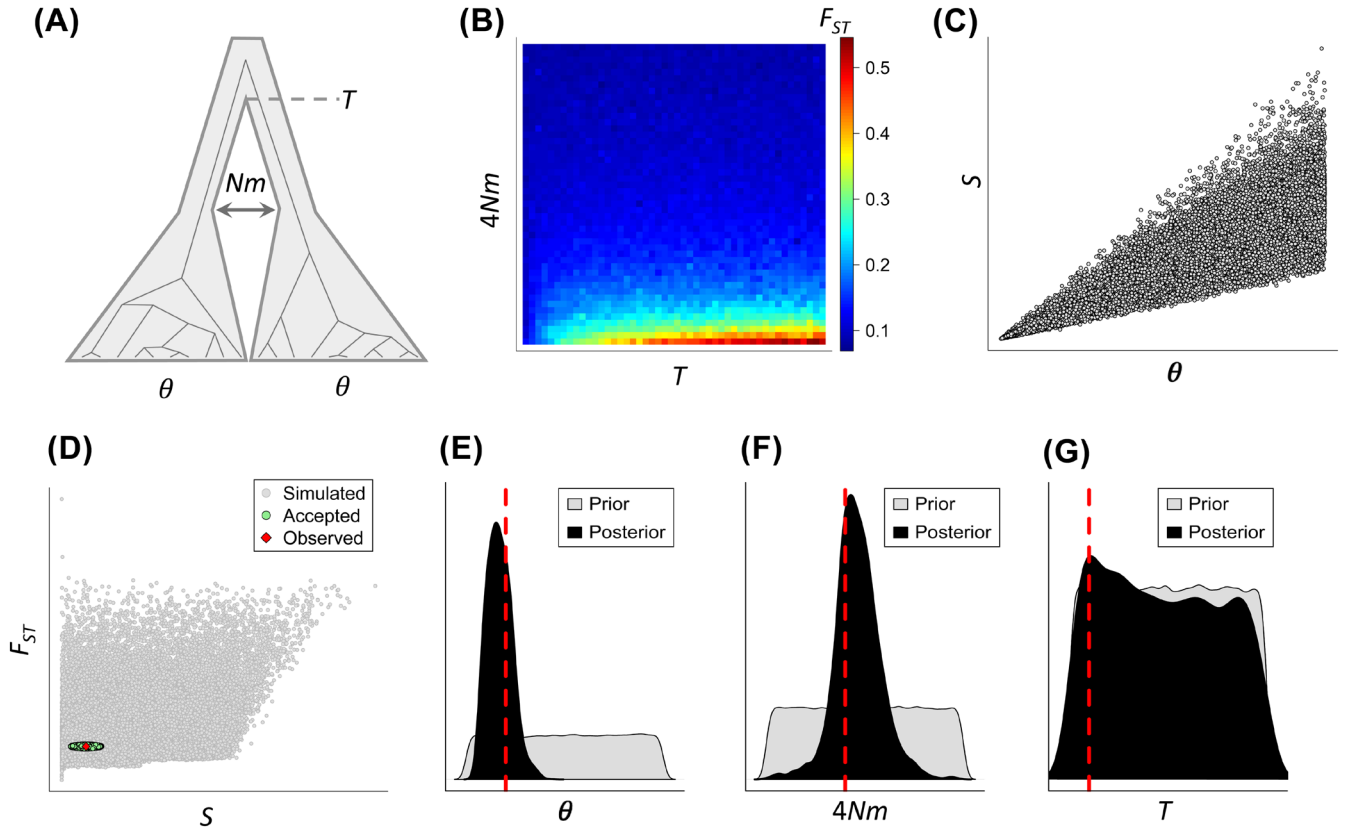


Figure 6. (A) A simple three-parameter model of population divergence, simulated 100 000 times using the R ‘coala’ package (Staab and Metzler 2016). Time elapses from top to bottom; ‘T’ signifies the time of divergence. Light grey shows population divergence while dark gray lines track lineage divergence. Nm is migration rate between lineages. (B–C) summary statistics from simulated datasets convey information about simulated parameter values (B – summary statistic F_{ST} is a function of both migration rate and divergence time, C – summary statistic total number of segregating sites S is a function of the parameter θ , which is defined as $4 \times$ effective population size \times mutation rate). (D) Simulated datasets summarized by F_{ST} and the number of segregating sites, grey dots correspond to datasets with Euclidean distances from the observed data (red diamond) above the 2.5th percentile, while green dots correspond to accepted simulations ($n=2500$). (E–G) Posterior distributions for model parameters, with simulated (known) values shown as red vertical dashed lines (E – θ , F – migration rate, $4Nm$, G – divergence time in units of $4N$ generations).

Second, it is essential to account for the uncertainty in data and processes used to inform the components in the ABC framework, specifically uncertainty in the habitat suitability matrix. The habitat matrix depends on choice of climate model, ENM algorithm that relates climate to occurrence (number of variables, type of algorithm and its parameterization), and transformation of habitat suitability into demographic parameters (Brown and Knowles 2012). Each choice results in a different habitat matrix (i.e. surface), but in the state-of-the-art typically only a single representative or mean habitat surface is used in the final analysis (Brown et al. 2016). However, different candidate algorithms with alternate parameterizations can produce qualitatively and quantitatively different predictions of habitat suitability (Warren and Siefert 2011, Radosavljevic and Anderson 2014, Qiao et al. 2015). Thus, using only a single habitat surface from a chosen model or an ensemble of models ignores uncertainty in this crucial step, leading to posterior credible intervals that are smaller than warranted (i.e. higher confidence than we should have).

The mapping between ENM-derived habitat suitability and abundance also requires better characterization, as available evidence suggests conflicting relationships (VanDerWal et al. 2009, Weber et al. 2017, Dallas and Hastings 2018). It is possible that this relationship varies by context and species, and thus should be incorporated into an integrative framework (cf. Brown and Knowles 2012). Of course, simulating demography with multiple algorithms, combinations of environmental layers, and mappings of habitat suitability to abundance rapidly increases the number of simulations that must be performed – a computational challenge which will be revisited in section 4.

Third, the summary statistics used in the ABC framework to evaluate the distance between simulations and data (and thus to estimate parameters of biogeographic history) can be improved. Traditionally in ABC, simple summary statistics are used, e.g. genetic differentiation, gene diversity, number of alleles or haplotypes, number of segregating sites in each sampled population, and bottleneck indices. These

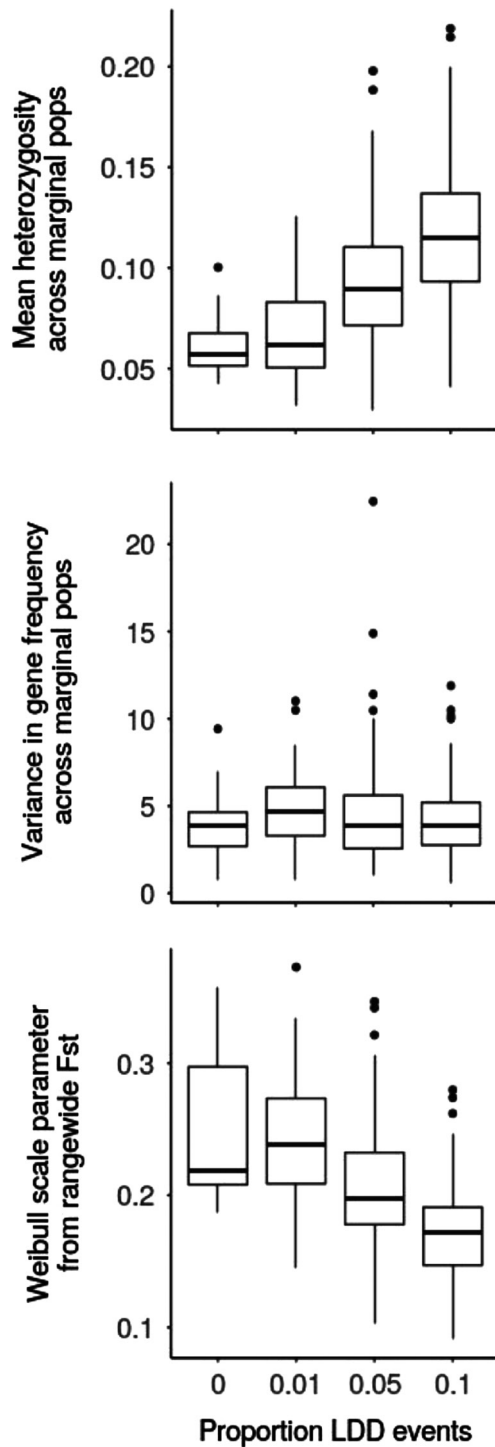


Figure 7. Summary of results from forward-time simulations illustrating how three spatial summary statistics respond to changes in a simulation input parameter. Simulations were conducted on a 2-dimensional landscape with a 10×10 grid of populations. Populations were colonized from a single refugium at the center of the bottom row and simulations proceeded until complete colonization of landscape. Genetic parameters are calculated from 100 co-dominant loci evolving through drift, dispersal and a mutation rate of 10^{-6} mutations/generation. In each panel the simulated

have often been applied in ABC studies of historical events (Benazzo et al. 2015). However, to make inference about spatial processes of postglacial colonization, expansion or contraction will require statistics that can capture spatial information for both single and multiple populations. Some new summary statistics have already been developed (Peter and Slatkin 2013, Dellicour et al. 2014, Alvarado-Serrano and Hickerson 2016, 2018); others have been proposed but not thoroughly tested (Kelleher et al. 2016, Prunier et al. 2017). In particular, spatial summary statistics may be especially useful when trying to make inference about spatial processes (e.g. long distance dispersal) or identify locations of refugia. This is an important advance on previous ABC work, which has mostly used non-spatial summary statistics such as mean pairwise F_{ST} or mean number of alleles. Based on our own exploratory simulations, we propose the following distance-dependent summary statistics (see Fig. 7 for other examples of within- and among-population diversity statistics): 1) the slope and intercept of a regression of pairwise divergence on Euclidean distances (isolation by distance), 2) the slope and intercept of regressions of latitude versus number of alleles, and 3) the breakpoint of a segmented regression of number of alleles versus latitude. Non-genetic summary statistics, i.e. any other description of states of the system that can be calculated on both observed and simulated data, have not been used frequently in ABC but could also prove useful. For example, when applying ABC to an individual-based model of white-starred robin *Pogonocichla stellata* population dynamics, Sirén et al. (2018) tested combinations of genetic data and information on the number, age and sex of captured birds as summary statistics. Their analyses highlight the utility of incorporating multiple data sources which provide complementary information for estimation of model parameters, into ABC analyses. Relevant non-genetic summary statistics for Quaternary migrations could include summaries of the entire species distribution (e.g. weighted geographic centroid or latitude of quantiles of species' abundance; Watts et al. 2013) and dynamic processes (e.g. age/stage structure).

Advance 2. Formal integration of fossil data into the inference framework

Maps of fossil occurrences are commonly used to define hypotheses about biogeographic history. However, to the best of our knowledge fossil occurrences have yet to be formally used in a quantitative approach that integrates genetic

Figure 7 Continued.

parameter of interest (and independent variable) is the proportion of long-distance dispersal events. Top panel: change in mean heterozygosity (gene diversity) in populations at top row of the 10×10 grid. Center panel: variance among populations in allele frequencies across the same top grid row as in the top panel. Bottom panel: best estimate of the scale parameter of a two-parameter Weibull probability density function fit to the range-wide distribution of among-population pairwise F_{ST} estimates.

and fossil data for inferring detailed biogeographic history. One approach to integrating fossil data is to combine it simultaneously with contemporary occurrences as inputs in an ecological niche model that leverages the temporal distribution of samples by calibrating the ENM using data from multiple time intervals simultaneously (Nogués-Bravo 2009, Maguire et al. 2016). A different approach, which is a natural extension to methods mentioned in section 2, is to allow fossil data to inform habitat suitability in an ABC framework. The abundance of pollen fossils in the sediments make them highly suitable for this use. Animal macrofossils may also be suitable for determining habitat suitability if there are a sufficient number of samples identifiable to species or genus (i.e. *Sigmodon*, Stangl and Dalquest 1991, or mussels, Popejoy et al. 2018). However, this approach is complicated by over- and under-representation of certain species in the fossil record. For plants, pollen vegetation models (PVMs) are used to reconstruct vegetation from fossil pollen records by modeling the processes that relate vegetation to observed pollen in the sediment, with the goal of accounting for these complications regarding representation. PVMs such as STEPPS (Dawson et al. 2016) can thus be used to create spatio-temporal maps of vegetation relative abundance, which could be used in place of an ENM as a proxy for habitat suitability. Alternatively, such maps could be merged with ENM-based maps using a weighting scheme to calculate habitat suitability (Fig. 5). Additionally, fossils that do not need interpretation through a process model (well-dated macrofossils, for example), can be integrated into ABC simulations as nodes which must be colonized by certain time periods for simulations to be retained for estimation.

Despite their strengths, neither of these two approaches would make use of information inherent in each data type (i.e. contemporary and fossil occurrences). The key innovation necessary will be to account for the diverse underlying processes generating each data type in one model. Although occurrence and fossil data both reflect habitat suitability, and to some degree abundance and successful dispersal (i.e. because populations of the species were present in that locale or region), a key difference is that fossil data additionally reflect site- and taxon-level differences in fossil deposition and formation (and, for pollen, production and transport). Another difference is that ENMs can reflect uninhabited yet suitable habitat, whereas fossils (if analyzed correctly) provide a sample from the realized distribution. The two are also measured differently, i.e. occurrence is usually unary or binary, while fossil data are either counts or proportions. Moreover, fossils may only be identifiable to the genus level, while contemporary occurrences are typically identifiable to species. An improvement on the above-mentioned efforts would employ, for example, a hierarchical Bayesian framework in which pollen dispersal and deposition surfaces for higher taxa are used as priors for species-specific ENMs. Alternatively a joint attribute model (Clark et al. 2017) for occurrence and pollen data could be used to integrate and faithfully reflect the characteristics of the data types. In sum, an analytical

synthesis would leverage the strengths in each type of data, i.e. most natural history occurrence data reflect contemporary distributions (i.e. the last ~250 years), while depositional pollen data reflect distribution across millennia.

Advance 3. Truly integrative models

Increasingly, the limiting assumptions of ‘traditional’ ENMs (namely, the species is at equilibrium with its environment and all inhabited locations are suitable) are being addressed by the development of a broad class of methods called dynamic range models (DRMs) which integrate multiple data types (Zurell et al. 2016). For example, Pagel and Schurr (2012) developed a Bayesian DRM composed of a sequence of conditional, probabilistic equations that describe abundance, detectability, dispersal, population growth, and the influence of the environment on carrying capacity. While powerful, existing DRMs have the following drawbacks: 1) they require detailed demographic data from many populations for multiple generations (impossible for most species), and 2) they ignore population genetic processes and information in other types of data (e.g. pollen). An alternative to a DRM is a hierarchical Bayesian metamodel that uses a mechanistic or correlative model to constrain an ENM using genetic, phenological, trait, pollen-vegetation, experimental, or other data (Talluto et al. 2016).

The next development would be a joint ENM/PVM/genetic range model where key processes that impact all of the data types are described and then formally linked to each set of observations. This will facilitate the joint estimation of state-space and process parameters because each data type is influenced in part by a set of common underlying demographic and genetic processes, and in part by unique and multi-scale processes. For example, genetic data not only provide information on migration rates and population size changes, but at fine scales can provide data on typical mating patterns and dispersal for a species. Similarly, environmental suitability estimated by a joint ENM/PVM could be used to inform reproduction and survival modeled in the demographic stage. By integrating the ENM, PVM, DRM and genetic components in a single inferential framework, the steep data requirements for DRMs may be obviated. For example, DRMs have been shown to perform poorly when long distance dispersal is common (Pagel and Schurr 2012), but genetic data may help constrain this weakly characterized yet critical aspect of demography. A joint inferential framework would also make it possible to incorporate additional information as available, such as field or experimental observations.

Advance 4. Modeling groups of species

Another innovation is to combine the described advances in a framework that incorporates communities, traits, and/or phylogenetics. The previous advances and examples in this review assumed that the goal was to understand biogeographic history for only a single focal species. A joint multi-species framework would allow species to ‘borrow’ inferential

strength from other species in the same geographic region (Chan et al. 2014). Also, traits facilitate colonization success (Feurdean et al. 2013, Lankau et al. 2015, Estrada et al. 2016), and so species with similar traits like dispersal mode, height, cold tolerance and successional status (or species that are closely related phylogenetically) may show similar timing and rate of colonization (see also Lawing and Matzke 2014). It is also possible that species within a community may have similar resource requirements and environmental tolerances, and therefore migrate in tandem. Of course, there is also evidence that many communities are not coherent through time (e.g. ‘no analog’ communities; Gill et al. 2009). Regardless, a Bayesian framework would allow integration of such prior knowledge, allowing multiple species to be analyzed jointly, using hyperparameters to model movement rates for similar species, and a single parameter that determines the ‘coherence’ of species’ responses. The idea of estimating community coherence is not new; in earlier work a coherence parameter was used to test hypotheses about the timing of species’ divergence (Hickerson et al. 2006, Stone et al. 2012) and simultaneous demographic expansion (Chan et al. 2014). The hyperparameter in these studies was the degree of synchronicity in timing of species’ divergence or range expansion, while the effective population sizes of each species were treated as nuisance parameters. As opposed to performing independent ABC studies for each species, multi-species hierarchical ABC (hABC) can utilize hyperparameters describing community level responses which link species-responses together (Hickerson et al. 2006). This approach facilitates the ability to quantitatively test hypotheses about simultaneous range expansions and can, in theory, be used to test for other shared parameters like proportion of long distance dispersal, carrying capacity and location of glacial refugia.

Section 4. Continued challenges and frontiers

This section presents challenges and frontiers for integrative, quantitative biogeographic inference. These challenges and frontiers include method validation, computational feasibility, sampling, new data sources, applications to taxa with sparse information in the fossil record, and genetic adaptation during range expansion.

The current state-of-the-art and future improvements will need to be fully validated for accuracy of parameter inference. For example, it would be important to assess uncertainty by making ABC inferences using different sets of components, (genetic+ENM, genetic+pollen and genetic+pollen+ENM), to understand the degree to which each data type influences posterior distributions. It will also be critical to assess sensitivity to the scaling of habitat suitability matrix with abundance (Brown and Knowles 2012, Brown et al. 2016), such as allowing the ENM to have strong or weak influence on demography. As noted above, a benefit of ABC is the ability to quantify confidence in model selection and parameter estimation with minimal additional computation (Bertorelle et al. 2010). To this end, datasets simulated under

candidate models (e.g. refugium location) or alternative parameter values are subjected to ABC model selection and parameter estimation analyses in leave-one-out cross validations. This is a standard practice for testing new methods in evolutionary genetics when we cannot observe the real state of the system (e.g. large scale ecosystem dynamics since the last glacial maximum) but we can observe such a system in silico (Bertorelle et al. 2010, Hoban 2014, Robinson et al. 2014, Zurell et al. 2016). Validation provides an assessment of model identifiability, expected accuracy of parameter estimates, and coverage of credible intervals.

ABC and range models typically already require cluster computers with hundreds of processors. More complex models, as we have outlined, will only increase computational demands and thus will require efficient algorithms, potentially exploiting computer architectures such as graphical processing units for highly parallel problems. Careful consideration of simulation details such as spatial and temporal resolution and extent are also important; there are tradeoffs in the ability of a model to resolve fine scale versus domain size processes. Scaling of population sizes in the demographic simulations may help to some degree (He et al. 2017). Also, even in comparatively simple coalescent models, summary statistic calculation can be computationally expensive relative to the coalescent simulation itself. To help avoid this expense, ‘checks’ based on simple (and thus computationally fast) summary statistics could be calculated first to determine if simulations should be retained, and if so then more demanding summary statistics could be calculated; to our knowledge such an approach has not been tested. Another way to reduce compute time is to reduce parameter space, possibly by first simulating a limited set of key processes using a general simulation model (Gotelli et al. 2009), then choosing likely parameter combinations via semi-manual tuning (Barnes and Clark 2017), and on this basis develop detailed simulations. Lastly, machine learning approaches are now helping to find better matches between observed and simulated data, thereby reducing computation needs by one to two orders of magnitude in recent studies (Blum and François 2010, Pudlo et al. 2016).

Future advances should ideally be open-source (to facilitate community contributions and further encourage broad application of integrative modeling approaches) while leveraging existing software and databases. For instance, existing R packages that access the GBIF (GBIF 2018) database (rgbif, Chamberlain and Boettiger 2017) and construct ecological niche models (ENMTools, Warren et al. 2010, wallace, Kass et al. 2018, enmSdm, Smith 2019) could be linked with packages for realistic models of population demography (e.g. rmetasim, Strand 2002, Strand and Niehaus 2016), coalescent simulations (e.g. coala, Staab and Metzler 2016), and analysis of population genetic data (e.g. strataG, Archer et al. 2016, poppr, Kamvar et al. 2014).

Computational power is not the only limitation. In spatial genetics and spatial ecology it is important to sample in the right spatial locations with enough samples to capture the

signal (Storfer et al. 2007). Bias in data collection can lead to erroneous conclusions (Phillips et al. 2009, Kalinowski 2011). As datasets are now trending towards more molecular markers and fewer individuals, spatially optimized sampling of individuals (i.e. choosing between transect or grid sampling) may become more important. Similar questions regard the number of observations, spatial and environmental coverage, and detectability for ENMs (Loiselle et al. 2008). Best practice recommendations for all data types are sorely needed. Fortunately, the same simulations that are used for validation can be used to determine optimal sampling strategies (Benazzo et al. 2015).

Excitingly, additional sources of data are becoming available. Ancient DNA has increased in reliability and information content such that intraspecific variation can be detected and simple genetic analyses performed (Palkopoulou et al. 2015). However, obtaining such data is still impossible for most research groups due to the facilities and expertise required, and protocols still need optimization. Moreover, as yet, ancient DNA analysis is usually only able to be performed on plastid markers (chloroplast, mitochondria) because there are many copies per cell. Still, population genetic analysis on ancient DNA is on the horizon (Llamas et al. 2016, Wagner et al. 2018). Another source of data is sediment cores from near-offshore locations (de Vernal and Hillaire-Marcel 2008, Leroy et al. 2013), which can reveal vegetation composition of land that was exposed during the Ice Age (possibly supporting refugial populations) but is now submerged.

As noted in the introduction, the modeling framework and future directions introduced here are particularly well-suited to application in plants, given the rich information in sedimentary pollen fossils. Extension of the integrative models to species with less representation in the fossil record will pose a significant challenge. In particular, relative abundance across space and time will be more challenging to estimate. However, even sparse fossil data may be usefully integrated as an indication of known presence in a region at a point in time. Such information could be used to constrain simulation histories or, potentially, used as a non-genetic summary statistic. For example, some animal taxa which were used by indigenous peoples may be common in assemblages from archeological sites, though models may be needed to account for the preservation of such remains. Also, ancient DNA, which is well preserved in some animal fossils, may compensate for fossil rarity by providing richer genetic data. Database developments in the future, including the expansion of existing fossil (e.g. Neotoma; Williams et al. 2018; PBDB; pbdb.org) and contemporary occurrence (e.g. GBIF; GBIF 2018) databases, will undoubtedly facilitate integration of these data in the types of modeling efforts proposed above.

Continued improvements in ENMs are occurring, such as use of non-point occurrence data like range maps (Merow et al. 2017), multi-species modeling (Nieto-Lugilde et al. 2018, Zhang et al. 2018), and accommodations for imperfect detection (Koshikana et al. 2017). Another interesting

avenue is to incorporate multiple molecular marker types and genomic-scale data. As yet, most genetic ABC studies use only one marker type (i.e. nuclear microsatellites). Plastid and nuclear genomes hold complementary information due to different modes of dispersal, effective population size and mutation rates (Petit et al. 2005). Additionally, microsatellites, large SNP datasets such as produced by RADseq, and sequences (from nuclear genomes and from chloroplasts) each offer different information content due to mutation and recombination, thus combining marker types should increase the ability of integrative modeling to make demographic inference, a point which to our knowledge is rarely exploited in the literature. Continuous declines in the cost of DNA sequencing, and new sequence capture approaches (e.g. Rapture, Ali et al. 2016), allow for cost-effective collection of genomic-scale data. These datasets offer increased power for estimating demographic parameters (e.g. contemporary effective population size; Waples et al. 2016). Likewise, a better understanding of how functional traits relate to the environment (Kostikova et al. 2013, Soudzilovskaia et al. 2013, Smith et al. 2017) coupled with growing databases on traits (Kattge et al. 2011) could help constrain models of species' spread (Angert et al. 2011).

It is important to note that the methods and models we have discussed all ignore adaptation via natural selection as well as genetic load and inbreeding depression. Thus, it is assumed that a species niche and fundamental traits such as dispersal do not change over time. However, species are already adapting to modern climate change, and adaptive evolutionary change may have occurred over the Quaternary. As one example, dispersal abilities can increase on the edge of a colonization front (Phillips et al. 2010). It may be infeasible to incorporate species-specific models of environmental adaptation into the simulations we describe, as detailed knowledge of the number of genes, their effect sizes and their selection coefficients exists for very few species (Tiffin and Ross-Ibarra 2014, Hoban et al. 2016, Csilléry et al. 2018). However, simple models of selection at a small number of genes (Schiffers et al. 2013) could be integrated into the framework outlined above, possibly making inferences more robust. Mathematical models of genetic variance and adaptation could also be useful (Polechová and Barton 2015). Alternatively, a parameter that allows a degree of niche shift could be used to allow the environmental relationship captured by the ENM (and thus the correlation between a given climate variable and probability of occurrence) to evolve (Kearney et al. 2009, Bush et al. 2016).

Summary and outlook

This paper is meant to serve as a guide for recent and future developments to the methods for quantitative inference of biogeographic history, along with advantages, disadvantages and challenges. We are not the first to highlight the need for a more robust integrative framework; as noted in

the Introduction, for decades discrepancies among expansion rates have suggested this need. Here, we propose a set of scaffolded and feasible advancements to existing methods. These advances increasingly integrate the three main lines of evidence (genetic, ENMs and fossil) used for inferring species' recent biogeographic histories. As noted above, a coherent integration of data types may not always lead to final resolution, but rather in some cases could lead to greater uncertainty. We emphasize that this is still a major advancement, i.e. a more complete characterization of uncertainty based on all available evidence and understanding will result in less-biased conclusions. This statement holds true even if the bounds of certainty are not reduced. To paraphrase the philosopher Carverth Read, 'It is better to be roughly right than precisely wrong.' Full description of uncertainty will also help identify avenues of improvement in datasets or models in the future. If advances such as those reviewed here can be achieved and verified, investigators will not only be able to infer parameters about refugia and rates of migration for single species, but also will finally answer long-standing questions about the nature of refugia (i.e. how species persist in environmentally stressful times), community coherence, niche stability and long-distance dispersal.

Acknowledgements – The authors thank members of the Scribner Lab at Michigan State University and Elizabeth Hobson and Andrew Hipp, for helpful comments and feedback, and Cindy Johnson for help in manuscript preparation.

Funding – This work was supported in part by NSF award 1759759 to the authors. This work was assisted through participation in the Range Shifts in a Warming World Workshop at the National Institute for Mathematical and Biological Synthesis, sponsored by the National Science Foundation through NSF Award no. DBI-1300426, with additional support from The Univ. of Tennessee, Knoxville. ABS was supported by the Alan Graham Fund in Global Change. SH is supported by The Center for Tree Science at The Morton Arboretum, while AES received a Senior Fellowship from The Center for Tree Science. AD was supported through NSF Awards 1241874 and 1241851. AES was supported in part by NSF Awards 1355106 and 1357386.

Conflicts of interest – None known.

Author contributions – All authors conceived of and planned this study, and co-wrote the manuscript. After first author, authors are listed alphabetically, recognizing equal author contributions.

References

- Ali, O. A. et al. 2016. RAD capture (Rapture): flexible and efficient sequence-based genotyping. – *Genetics* 202: 389–400.
- Alvarado-Serrano, D. F. and Hickerson, M. J. 2016. Spatially explicit summary statistics for historical population genetic inference. – *Methods Ecol. Evol.* 7: 418–427.
- Alvarado-Serrano, D. F. and Knowles, L. L. 2014. Ecological niche models in phylogeographic studies: applications, advances and precautions. – *Mol. Ecol. Resour.* 14: 233–248.
- Alvarado-Serrano, D. F. and Hickerson, M. J. 2018. Detecting spatial dynamics of range expansions with geo-referenced genome-wide SNP data and the geographic spectrum of shared alleles. – *bioRxiv* 457556. doi: <https://doi.org/10.1101/457556>.
- Anderson, R. P. 2017. When and how should biotic interactions be considered in models of species niches and distributions? – *J. Biogeogr.* 44: 8–17.
- Antoniazza, S. et al. 2014. Natural selection in a postglacial range expansion: the case of the colour cline in the European barn owl. – *Mol. Ecol.* 23: 5508–5523.
- Archer, F. I. et al. 2016. strataG: an R package for manipulating, summarizing and analysing population genetic data. – *Mol. Ecol. Resour.* 17: 5–11.
- Angert, A. L. et al. 2011. Do species' traits predict recent shifts at expanding range edges? – *Ecol. Lett.* 14: 677–689.
- Arenas, M. et al. 2012. Consequences of range contractions and range shifts on molecular diversity. – *Mol. Biol. Evol.* 29: 207–218.
- Avise, J. C. et al. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. – *Annu. Rev. Ecol. Syst.* 18: 489–522.
- Bagley, J. C. et al. 2018. Genotyping-by-sequencing and ecological niche modeling illuminate phylogeography, admixture and Pleistocene range dynamics in quaking aspen (*Populus tremuloides*). – *PeerJ* 6: e27162v1. DOI: 10.7287/peerj.preprints.27162v1.
- Barnes, R. and Clark, A. T. 2017. Sixty-five million years of change in temperature and topography explain evolutionary history in eastern North American plethodontid salamanders. – *Am. Nat.* 190: E1–E12.
- Barnosky, A. D. et al. 2017. Merging paleobiology with conservation biology to guide the future of terrestrial ecosystems. – *Science* 355: eaah4787.
- Barve, N. et al. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. – *Ecol. Model.* 222: 1810–1819.
- Beaumont, M. A. 2010. Approximate Bayesian computation in evolution and ecology. – *Annu. Rev. Ecol. Evol. Syst.* 41: 379–406.
- Beaumont, M. A. and Nichols, R. A. 1996. Evaluating loci for use in the genetic analysis of population structure. – *Proc. R. Soc. B* 263: 1619–1626.
- Benazzo, A. et al. 2015. Using ABC and microsatellite data to detect multiple introductions of invasive species from a single source. – *Heredity* 115: 262–272.
- Bertorelle, G. et al. 2010. ABC as a flexible framework to estimate demography over space and time: some cons, many pros. – *Mol. Ecol.* 19: 2609–2625.
- Bialozyt, R. et al. 2006. Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. – *J. Evol. Biol.* 19: 12–20.
- Blaauw, M. and Christen, J. A. 2011. Flexible paleoclimate age-depth models using an autoregressive gamma process. – *Bayesian Anal.* 6: 457–474.
- Blum, M. G. B. and François, O. 2010. Non-linear regression models for approximate Bayesian computation. – *Stat. Comput.* 20: 63–73.
- Brown, J. L. and Knowles, L. L. 2012. Spatially explicit models of dynamic histories: examination of the genetic consequences of Pleistocene glaciation and recent climate change on the American *Pika*. – *Mol. Ecol.* 21: 3757–3775.
- Brown, J. L. et al. 2016. Predicting the genetic consequences of future climate change: the power of coupling spatial demography,

- the coalescent and historical landscape changes. – *Am. J. Bot.* 103: 153–163.
- Bush, A. et al. 2016. Incorporating evolutionary adaptation in species distribution modelling reduces projected vulnerability to climate change. – *Ecol. Lett.* 19: 1468–1478.
- Campos, P. F. et al. 2010. Ancient DNA analyses exclude humans as the driving force behind late Pleistocene musk ox (*Ovibos moschatus*) population dynamics. – *Proc. Natl Acad. Sci. USA* 107: 5675–5680.
- Carnaval, A. C. and Moritz, C. 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. – *J. Biogeogr.* 35: 1187–1201.
- Carnaval, A. C. et al. 2009. Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. – *Science* 323: 785–789.
- Chamberlain, S. and Boettiger, C. 2017. R Python, and Ruby clients for GBIF species occurrence data. – *PeerJ* 5: e3304v1.
- Chan, Y. L. et al. 2014. Detecting concerted demographic response across community assemblages using hierarchical approximate Bayesian computation. – *Mol. Biol. Evol.* 31: 2501–2515.
- Cheddadi, R. et al. 2014. Revisiting tree-migration rates: *Abies alba* (Mill.), a case study. – *Veg. Hist. Archaeobot.* 23: 113–122.
- Clark, J. S. 1998. Why trees migrate so fast: confronting theory with dispersal biology and the paleorecord. – *Am. Nat.* 152: 204–224.
- Clark, J. S. et al. 2017. Generalized joint attribute modeling for biodiversity analysis: median-zero, multivariate, multifarious data. – *Ecol. Monogr.* 87: 34–56.
- Collevatti, R. G. et al. 2015a. Multi-model inference in comparative phylogeography: an integrative approach based on multiple lines of evidence. – *Front. Genet.* 6: 31.
- Collevatti, R. G. et al. 2015b. Evolutionary and population genetics relaxed random walk model coupled with ecological niche modeling unravel the dispersal dynamics of a Neotropical savanna tree species in the deeper Quaternary. – *Front. Plant Sci.* 6: 1–15.
- Csilléry, K. et al. 2010. Approximate Bayesian computation (ABC) in practice. – *Trends Ecol. Evol.* 25: 410–418.
- Csilléry, K. et al. 2018. Detecting the genomic signal of polygenic adaptation and the role of epistasis in evolution. – *Mol. Ecol.* 27: 606–612.
- Dallas, T. A. and Hastings, A. 2018. Habitat suitability estimated by niche models is largely unrelated to species abundance. – *Global Ecol. Biogeogr.* 27: 1448–1456.
- 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.
- Davis, M. B. 1963. On the theory of pollen analysis. – *Am. J. Sci.* 261: 897–912.
- Davis, M. B. 1983. Quaternary history of deciduous forests of eastern North America and Europe. – *Ann. Miss. Bot. Gard.* 70: 550–563.
- de Bruyn, M. et al. 2011. Faunal histories from Holocene ancient DNA. – *Trends Ecol. Evol.* 26: 405–413.
- de Lafontaine, G. et al. 2014. Beyond skepticism: uncovering cryptic refugia using multiple lines of evidence. – *New Phytol.* 204: 450–454.
- Dellicour, S. et al. 2014. Inferring the past and present connectivity across the range of a North American leaf beetle: combining ecological niche modeling and a geographically explicit model of coalescence. – *Evolution* 68: 2371–2385.
- de Vernal, A. and Hillaire-Marcel, C. 2008. Natural variability of Greenland climate, vegetation and ice volume during the past million years. – *Science* 320: 1622–1625.
- Drummond, A. J. and Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. – *BMC Evol. Biol.* 7: 214.
- Espíndola, A. et al. 2012. Predicting present and future intra-specific genetic structure through niche hindcasting across 24 millennia. – *Ecol. Lett.* 15: 649–657.
- Estoup, A. et al. 2001. Inferring population history from microsatellite and enzyme data in serially introduced cane toads, *Bufo marinus*. – *Genetics* 159: 1671–1687.
- Estoup, A. et al. 2010. Combining genetic, historical and geographical data to reconstruct the dynamics of bioinvasions: application to the cane toad *Bufo marinus*. – *Mol. Ecol. Res.* 10: 886–901.
- Estrada, A. et al. 2016. Usefulness of species traits in predicting range shifts. – *Trends Ecol. Evol.* 31: 190–203.
- Fady, B. et al. 2016. Forests and global change: what can genetics contribute to the major forest management and policy challenges of the twenty-first century? – *Reg. Environ. Change* 16: 927–939.
- Feurdean, A. et al. 2013. 12 000-years of fire regime drivers in the lowlands of Transylvania (central-eastern Europe): a data-model approach. – *Q. Sci. Rev.* 81: 8–61.
- Ferris, C. et al. 1993. Native oak chloroplasts reveal an ancient divide across Europe. – *Mol. Ecol.* 2: 337–343.
- 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.
- Fordham, D. A. et al. 2013. Tools for integrating range change, extinction risk and climate change information into conservation management. – *Ecography* 36: 956–964.
- 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.
- Forester, B. R. et al. 2013. Integrating ensemble species distribution modeling and statistical phylogeography to inform projections of climate change impacts on species distributions. – *Divers. Distrib.* 19: 1480–1495.
- François, O. et al. 2008. Demographic history of European populations of *Arabidopsis thaliana*. – *PLoS Genet.* 4: e1000075.
- Franklin, J. 2010. Mapping species distributions: spatial inference and prediction, 1st edn. – Cambridge Univ. Press.
- Franklin, J. et al. 2013. Modeling plant species distributions under future climates: how fine do climate projections need to be? – *Global Change Biol.* 19: 473–483.
- Fuentes-Utrilla, P. et al. 2014. Extending glacial refugia for a European tree: genetic markers show that Iberian populations of white elm are native relicts and not introductions. – *Heredity* 112: 105–113.
- Gavin, D. G. et al. 2014. Climate refugia: joint inference from fossil records, species distribution models and phylogeography. – *New Phytol.* 204: 7–54.
- Gill, J. L. et al. 2009. Pleistocene megafaunal collapse, novel plant communities and enhanced fire regimes in North America. – *Science* 326: 1100–1103.
- GBIF: The Global Biodiversity Information Facility 2018. What is GBIF? – <www.gbif.org/what-is-gbif>, accessed 12 September 2018.
- Gong, W. et al. 2008. Phylogeography of a living fossil: Pleistocene glaciations forced *Ginkgo biloba* L. (Ginkgoaceae) into two

- refuge areas in China with limited subsequent postglacial expansion. – *Mol. Phylogenet. Evol.* 48: 1094–1105.
- Gotelli, N. J. et al. 2009. Patterns and causes of species richness: a general simulation model for macroecology. – *Ecol. Lett.* 12: 873–886.
- Grivet, D. et al. 2006. Contrasting patterns of historical colonization in white oaks (*Quercus* spp.) in California and Europe. – *Mol. Ecol.* 15: 4085–4093.
- Guillera-Aroita, G. et al. 2015. Is my species distribution model fit for purpose? Matching data and models to applications. – *Global Ecol. Biogeogr.* 24: 276–292.
- Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – *Ecol. Lett.* 8: 993–1009.
- Guralnick, R. 2007. Differential effects of past climate warming on mountain and flatland species distributions: a multispecies North American mammal assessment. – *Global Ecol. Biogeogr.* 16: 14–23.
- He, Q. et al. 2013. Integrative testing of how environments from the past to the present shape genetic structure across landscapes. – *Evolution* 67: 3386–3402.
- He, Q. et al. 2017. Inferring the geographic origin of a range expansion: latitudinal and longitudinal coordinates inferred from genomic data in an ABC framework with the program X-ORIGIN. – *Mol. Ecol.* 12: 3218–3221.
- Hefley, T. J. et al. 2017. Bias correction of bounded location errors in presence-only data. – *Methods Ecol. Evol.* 8: 1566–1573.
- Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. – *Nature* 405: 907.
- Hickerson, M. J. et al. 2006. Test for simultaneous divergence using approximate Bayesian computation. – *Evolution* 60: 2435–53.
- Hoban, S. 2014. An overview of the utility of population simulation software in molecular ecology. – *Mol. Ecol.* 23: 2383–401.
- Hoban, S. et al. 2012. Computer simulations: tools for population and evolutionary genetics. – *Nat. Rev. Genet.* 13: 110–22.
- Hoban, S. et al. 2016. Finding the genomic basis of local adaptation: pitfalls, practical solutions and future directions. – *Am. Nat.* 188: 379–397.
- Jackson, S. T. and Overpeck, J. T. 2000. Responses of plant populations and communities to environmental changes of the late Quaternary. – *Paleobiology* 26: 194–200.
- Jackson, N. D. et al. 2017. PHRAPL: phylogeographic inference using approximate likelihoods. – *Syst. Biol.* 66: 1045–1053.
- Kalinowski, S. T. 2011. The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. – *Heredity* 106: 625–632.
- Kamvar, Z. N. et al. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. – *PeerJ* 2: e281.
- Karger, D. N. et al. 2017. Climatologies at high resolution for the Earth's land surface areas. – *Sci. Data* 4: 170122.
- Kass, J. M. et al. 2018. Wallace: a flexible platform for reproducible modeling of species niches and distributions built for community expansion. – *Methods Ecol. Evol.* 9: 1151–1156.
- Kattge, J. et al. 2011. TRY – a global database of plant traits. – *Global Change Biol.* 17: 2905–2935.
- Kearney, M. et al. 2009. Integrating biophysical models and evolutionary theory to predict climatic impacts on species' ranges: the dengue mosquito *Aedes aegypti* in Australia. – *Funct. Ecol.* 23: 528–538.
- Kelleher, J. et al. 2016. Spread of pedigree versus genetic ancestry in spatially distributed populations. – *Theor. Pop. Biol.* 108: 1–12.
- Knowles, L. L. and Alvarado-Serrano, D. F. 2010. Exploring the population genetic consequences of the colonization process with spatio-temporally explicit models: Insights from coupled ecological, demographic and genetic models in montane grasshoppers. – *Mol. Ecol.* 19: 3727–3745.
- Koshikana, V. et al. 2017. Integrated species distribution models: combining presence-background data and site-occupancy data with imperfect detection. – *Methods Ecol. Evol.* 8: 420–430.
- Kostikova, A. et al. 2013. Linking life-history traits, ecology and niche breadth evolution in North American eriogonoids (Polygonaceae). – *Am. Nat.* 182: 760–774.
- Kujawa, E. R. et al. 2016. The effects of anthropogenic land cover change on pollen–vegetation relationships in the American Midwest. – *Anthropocene* 15: 60–71.
- Lankau, R. A. et al. 2015. Mycorrhizal strategies of tree species correlate with trailing range edge responses to current and past climate change. – *Ecology* 96: 1451–1458.
- Lawing, A. M. and Matzke, N. J. 2014. Conservation paleobiology needs phylogenetic methods. – *Ecography* 37: 1109–1122.
- Leroy, S. A. et al. 2013. Holocene vegetation history and sea level changes in the SE corner of the Caspian Sea: relevance to SW Asia climate. – *Q. Sci. Rev.* 70: 28–47.
- Llamas, B. et al. 2016. Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. – *Sci. Adv.* 2: e1501385.
- Loiselle, B. A. et al. 2008. Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? – *J. Biogeogr.* 35: 105–116.
- Lorenz, D. J. et al. 2016. Downscaled and debiased climate simulations for North America from 21 000 years ago to 2100AD. – *Sci. Data* 3: 160048.
- Lotterhos, K. E. and Whitlock, M. C. 2014. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. – *Mol. Ecol.* 23: 2178–2192.
- Lumibao, C. Y. et al. 2017. Ice ages leave genetic diversity 'hotspots' in Europe but not in Eastern North America. – *Ecol. Lett.* 20: 1459–1468.
- Luikart, G. et al. 1998. Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. – *Mol. Ecol.* 7: 963–974.
- 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.
- Maguire, K. C. et al. 2016. Controlled comparison of species- and community-level models across novel climates and communities. – *Phil. Trans. R. Soc. B* 283: 20152817.
- Marino, I. A. M. et al. 2013. Evidence for past and present hybridization in three Antarctic icefish species provides new perspectives on an evolutionary radiation. – *Mol. Ecol.* 22: 5148–5161.
- Massatti, R. and Knowles, L. L. 2016. Contrasting support for alternative models of genomic variation based on microhabitat preference: species-specific effects of climate change in alpine sedges. – *Mol. Ecol.* 25: 3974–3986.
- McLachlan, J. S. and Clark, J. S. 2004. Reconstructing historical ranges with fossil data at continental scales. – *For. Ecol. Manage.* 197: 39–147.
- McLachlan, J. S. et al. 2005. Molecular indicators of tree migration capacity under rapid climate change. – *Ecology* 86: 2088–2098.

- Merow, C. et al. 2017. Integrating occurrence data and expert maps for improved species range predictions. – *Global Ecol. Biogeogr.* 26: 243–258.
- Midgley, G. F. et al. 2010. BioMove – an integrated platform simulating the dynamic response of species to environmental change. – *Ecography* 33: 612–616.
- Milligan, B. G. et al. 1994. Conservation genetics: beyond the maintenance of marker diversity. – *Mol. Ecol.* 3: 423–435.
- Mona, S. et al. 2014. Genetic consequences of habitat fragmentation during a range expansion. – *Heredity* 112: 291–299.
- Nielsen, R. and Wakeley, J. 2001. Distinguishing migration from isolation: a Markov chain Monte Carlo approach. – *Genetics* 158: 885–896.
- Nieto-Lugilde, D. et al. 2018. Multiresponse algorithms for community-level modelling: review of theory, applications and comparison to species distribution models. – *Methods Ecol. Evol.* 9: 834–848.
- Nogués-Bravo, D. 2009. Predicting the past distribution of species climatic niches. – *Global Ecol. Biogeogr.* 18: 521–531.
- Nogués-Bravo, D. et al. 2018. Cracking the code of biodiversity responses to past climate change. – *Trends Ecol. Evol.* 33: 765–776.
- Ovaskainen, O. et al. 2010. Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. – *Ecology* 91: 2514–2521.
- Overpeck, J. T. et al. 1985. Quantitative interpretation of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs. – *Q. Res.* 23: 87–108.
- Pagel, J. and Schurr, F. M. 2012. Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. – *Global Ecol. Biogeogr.* 21: 293–304.
- Palkopoulou, E. et al. 2015. Complete genomes reveal signatures of demographic and genetic declines in the woolly mammoth. – *Curr. Biol.* 25: 1395–1400.
- Parducci, L. et al. 2012. Glacial survival of boreal trees in northern Scandinavia. – *Science* 335: 1083–1086.
- Parnell, A. C. et al. 2011. A review of statistical chronology models for high-resolution, proxy-based Holocene palaeoenvironmental reconstruction. – *Q. Sci. Rev.* 30: 2948–2960.
- Parsons, R. W. and Prentice, I. C. 1981. Statistical approaches to R-values and the pollen–vegetation relationship. – *Rev. Palaeobot. Palynol.* 32: 127–152.
- Patiño, J. et al. 2015. Approximate Bayesian computation reveals the crucial role of oceanic islands for the assembly of continental biodiversity. – *Syst. Biol.* 64: 579–589.
- Peter, B. M. and Slatkin, M. 2013. Detecting range expansions from genetic data. – *Evolution* 67: 3274–3289.
- Peterson, A. T. et al. 2011. Ecological niches and geographic distributions. *Monogr. Popul. Biol.* 49. – Princeton Univ. Press.
- Petit, R. J. et al. 1997. Chloroplast DNA footprints of postglacial recolonization by oaks. – *Proc. Natl Acad. Sci. USA* 94: 9996–10001.
- Petit, R. J. et al. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. – *Science* 300: 1563–1565.
- Petit, R. J. et al. 2005. Invited review: comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. – *Mol. Ecol.* 14: 689–701.
- Petit, R. J. et al. 2008. Forests of the past: a window to future changes. – *Science* 320: 1450–1452.
- Phillips, S. J. et al. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. – *Ecol. Appl.* 19: 181–197.
- Phillips, B. L. et al. 2010. Evolutionarily accelerated invasions: the rate of dispersal evolves upwards during the range advance of cane toads. – *J. Evol. Biol.* 23: 2595–2601.
- Puong, M. A. et al. 2014. Delimiting species in the genus *Otospermophilus* (Rodentia: Sciuridae), using genetics, ecology and morphology. – *Biol. J. Linn. Soc.* 113: 1136–1151.
- Polechová, J. and Barton, N. H. 2015. Limits to adaptation along environmental gradients. – *Proc. Natl Acad. Sci. USA* 112: 6401–6406.
- Popejoy, T. et al. 2018. Prioritizing sites for conservation based on similarity to historical baselines and feasibility of protection. – *Conserv. Biol.* 32: 1118–1127.
- Prates, I. et al. 2016. Inferring responses to climate dynamics from historical demography in neotropical forest lizards. – *Proc. Natl Acad. Sci. USA* 113: 7978–7985.
- Pritchard, J. K. et al. 1999. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. – *Mol. Biol. Evol.* 16: 1791–1798.
- Prunier, J. G. et al. 2017. Regression commonality analyses on hierarchical genetic distances. – *Ecography* 40: 1412–1425.
- Pudlo, P. et al. 2016. Reliable ABC model choice via random forests. – *Bioinformatics* 32: 859–866.
- Qiao, H. et al. 2015. No silver bullets in correlative ecological niche modeling: Insights from testing among many potential algorithms for niche estimation. – *Methods Ecol. Evol.* 6: 1126–1136.
- Radosavljevic, A. and Anderson, R. P. 2014. Making better Maxent models of species distributions: complexity, overfitting and evaluation. – *J. Biogeogr.* 41: 629–643.
- Ray, N. et al. 2010. SPLATCHE2: a spatially explicit simulation framework for complex demography, genetic admixture and recombination. – *Bioinformatics* 26: 2993–2994.
- Reid, C. 1899. The origin of the British flora. – Dulau & Company.
- Richards, C. L. et al. 2007. Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. – *J. Biogeogr.* 34: 1833–1845.
- Roberts, D. R. and Hamann, A. 2012. Predicting potential climate change impacts with bioclimate envelope models: a palaeoecological perspective. – *Global Ecol. Biogeogr.* 21: 121–133.
- Roberts, D. R. and Hamann, A. 2015. Glacial refugia and modern genetic diversity of 22 western North American tree species. – *Proc. R. Soc. B* 282: 20142903.
- Robinson, J. D. et al. 2014. ABC inference of multi-population divergence with admixture from unphased population genomic data. – *Mol. Ecol.* 23: 4458–4471.
- Rougemont, Q. and Bernatchez, L. 2018. The demographic history of Atlantic salmon (*Salmo salar*) across its distribution range reconstructed from approximate Bayesian computations*. – *Evolution* 72: 1261–1277.
- Santos, M. J. et al. 2017. The relative influence of change in habitat and climate on elevation range limits in small mammals in Yosemite National Park, California, U.S.A.. – *Clim. Change Resp.* 4: 7.
- Schaub, M. and Abadi, F. 2011. Integrated population models: a novel analysis framework for deeper insights into population dynamics. – *J. Ornithol.* 152: 227–237.
- Schiffers, K. et al. 2013. Limited evolutionary rescue of locally adapted populations facing climate change. – *Phil. Trans. R. Soc. B* 368: 0083.

- Shafer, A. B. A. et al. 2015. Demographic inferences using short-read genomic data in an approximate Bayesian computation framework: in silico evaluation of power, biases and proof of concept in Atlantic walrus. – *Mol. Ecol.* 24: 328–345.
- Sirén, J. et al. 2018. Assessing the dynamics of natural populations by fitting individual-based models with approximate Bayesian computation. – *Methods Ecol. Evol.* 2018: 1286–1295.
- Smith, A. B. 2013. The relative influence of temperature, moisture and their interaction on range limits of mammals over the past century. – *Global Ecol. Biogeogr.* 22: 334–343.
- Smith, A. B. 2019. *enmSdm*. GitHub repository. – <www.github.com/adamlilith/enmSdm>.
- Smith, A. B. et al. 2017. Phenotypic distribution models corroborate species distribution models: a shift in the role and prevalence of a dominant prairie grass in response to climate change. – *Global Change Biol.* 23: 4365–4375.
- Smith, A. B. et al. 2019. Niche estimation above and below the species level. – *Trends Ecol. Evol.* 34: 260–273.
- Soudzilovskaia, N. et al. 2013. Functional traits predict relationship between plant abundance dynamic and long-term climate warming. – *Proc. Natl Acad. Sci. USA* 110: 18180–18184.
- Staab, P. R. and Metzler, D. 2016. Coala: an R framework for coalescent simulation. – *Bioinformatics* 32: 1903–1904.
- Stangl, F. and Dalquest, W. W. 1991. Historical biogeography of *Sigmodon-Ochrognathus* in Texas. – *Tex. J. Sci.* 43: 127–131.
- Stone, G. N. et al. 2012. Reconstructing community assembly in time and space reveals enemy escape in a Western Palearctic insect community. – *Curr. Biol.* 22: 532–537.
- Storfer, A. et al. 2007. Putting the ‘landscape’ in landscape genetics. – *Heredity* 98: 128–142.
- Strand, A. 2002. Metasim 1.0: an individual-based environment for simulating population genetics of complex population dynamics. – *Mol. Ecol. Notes* 2: 373–376.
- Strand, A. and Niehaus, J. 2016. rmetasim: an individual-based population genetic simulation environment. R package ver. 3.1.0. – <<https://github.com/stranda/rmetasim>>.
- Sugita, S. 2007a. Theory of quantitative reconstruction of vegetation I: pollen from large sites REVEALS regional vegetation composition. – *Holocene* 17: 229–241.
- Sugita, S. 2007b. Theory of quantitative reconstruction of vegetation II: all you need is LOVE. – *Holocene* 17: 243–257.
- Sukumaran, J. and Knowles, L. L. 2018. Trait-dependent biogeography: (Re) integrating biology into probabilistic historical biogeographical models. – *Trends Ecol. Evol.* 33: 390–398.
- Svenning, J.-C. et al. 2011. Applications of species distribution modeling to paleobiology. – *Q. Sci. Rev.* 30: 2930–2947.
- Talluto, M. V. et al. 2016. Cross-scale integration of knowledge for predicting species ranges: a metamodeling framework. – *Global Ecol. Biogeogr.* 25: 238–249.
- Tavare, S. et al. 1997. Inferring coalescence times from DNA sequence data. – *Genetics* 145: 505–518.
- Thornton, K. and Andolfatto, P. 2006. Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of *Drosophila melanogaster*. – *Genetics* 172: 1607–1619.
- Thuiller, W. et al. 2013. A road map for integrating eco-evolutionary processes into biodiversity models. – *Ecol. Lett.* 16: 94–105.
- Tiffin, P. and Ross-Ibarra, J. 2014. Advances and limits of using population genetics to understand local adaptation. – *Trends Ecol. Evol.* 29: 673–680.
- Valladares, F. et al. 2014. The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. – *Ecol. Lett.* 17: 1351–1364.
- Varela, S. et al. 2015. A short guide to the climatic variables of the last glacial maximum for biogeographers. – *PLoS One* 10: e0129037.
- Velasquez-Tibata et al. 2016 Using measurement error models to account for georeferencing error in species distribution models. – *Ecography* 39: 305–316.
- VanDerWal, J. et al. 2009. Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. – *Am. Nat.* 174: 282–291.
- Wagner, S. et al. 2018. High-throughput DNA sequencing of ancient wood. – *Mol. Ecol.* 27: 1138–1154.
- Waples, R. K. et al. 2016. Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. – *Heredity* 117: 233.
- Warren, D. L. and Siefert, S. N. 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. – *Ecol. Appl.* 21: 335–342.
- Warren, D. L. et al. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. – *Ecography* 33: 607–611.
- Watts, M. J. et al. 2013. Tracking shifting range margins using geographical centroids of metapopulations weighted by population density. – *Ecol. Model.* 269: 61–69.
- Weber, M. M. et al. 2017. Is there a correlation between abundance and environmental suitability derived from ecological niche modeling? A meta-analysis. – *Ecography* 40: 817–828.
- White, T. A. et al. 2013 Adaptive evolution during an ongoing range expansion: the invasive bank vole (*Myodes glareolus*) in Ireland. – *Mol. Ecol.* 22: 2971–2985.
- Wiens, J. A. et al. 2009. Niches, models and climate change: assessing the assumptions and uncertainties. – *Proc. Natl Acad. Sci. USA* 106: 19729–19736.
- Williams, J. W. et al. 2009. Rapid responses of the prairie-forest ecotone to early Holocene aridity in mid-continental North America. – *Global Planet. Change* 66: 195–207.
- Williams, J. W. et al. 2018. The Neotoma Paleoecology Database, a multiproxy, international, community-curated data resource. – *Q. Res.* 89: 156–177.
- Woolbright, S. A. et al. 2014. Climate relicts and their associated communities as natural ecology and evolution laboratories. – *Trends Ecol. Evol.* 29: 406–416.
- Zhang, C. et al. 2018. Comparing the prediction of joint species distribution models with respect to characteristics of sampling data. – *Ecography* 41: 1–12.
- Zhao, Y. et al. 2010. Hemlock (*Tsuga canadensis*) declines at 9800 and 5300 cal. yr BP caused by Holocene climatic shifts in northeastern North America. – *Holocene* 20: 877–886.
- Zurell, D. et al. 2016. Benchmarking novel approaches for modelling species range dynamics. – *Global Change Biol.* 22: 2651–2664.