An evolving view of phylogenetic biogeography

Michael J. Landis[1], **Joel L. Cracraft**[2], **Isabel Sanmartín**[3]

[1] *Department of Biology, Washington University, One Brookings Drive, St. Louis, Missouri 63130, USA*

[2] *Department of Ornithology, American Museum of Natural History, New York 10034, USA*

[3] *Real Jardín Botánico (RJB), CSIC, Plaza de Murillo 2, 28014 Madrid, Spain*

Email: michael.landis@wustl.edu

Running head: EVOLVING VIEW OF PHYLOGENETIC BIOGEOGRAPHY

Abstract: Biogeography is intrinsically linked to evolution as a process and to systematics as a practice. Phylogenetic biogeography, in particular, studies the distribution of life in space over time through the lens of common ancestry. Over the past century, new biological and geological discoveries, theoretical frameworks, and methodological techniques revolutionized how we understand why species have come to live where they do. This perspective piece orients readers to major advances, changes, and conflicts from the phylogenetic biogeography literature, much of which was centered around articles published in this very journal. As part of our survey, we also highlight areas that were historically active, remain biologically significant, and deserve renewed attention.

Keywords: biogeography; phylogenetics; cladistics; history; evolving view

All evolution occurs within a spatial context, making it crucial to consider where the ancestors of a species once lived to understand that species' present way of existence. To this end, historical biogeography studies how life evolved in space over time, principally in terms of the dispersal, extinction, and speciation processes that operate upon species, and in terms of the geological, climatological, and ecological processes that operate upon landscapes. Phylogenetic biogeography, the focus of this paper, extends the historical perspective to explicitly consider common ancestry among species. What information does a phylogenetic tree carry about the biogeographic history of its taxa and the landscape they evolved within? And, what do we learn when faced with biogeographic and phylogenetic patterns for a clade that are seemingly at odds with each other, defying the simplest historical explanation?

Dreaming big, a fully explanatory theory of phylogenetic biogeography would be able to perfectly reconstruct when and where each species ever lived, determine how those species originated, and measure how likely each was to move, speciate, evolve, or go extinct. This is aspirational, idealistic, and probably impossible. As in other non-experimental disciplines, phylogenetic biogeography must rely upon correlational evidence: we cannot observe evolution at work over deep timescales, so we need to infer past biogeographical events and ancestral distributions based on evidence collected at the present alone. Despite being unattainable in practice, phylogenetic biogeographers nonetheless wish to develop a predictive theory that is as accurate and useful as can be. How they have pursued this ideal has shaped how the field frames its research problems, solutions, and discourse.

This article focuses on three central methodological goals of phylogenetic biogeography: reconstruction of ancestral species ranges, inference of speciation mode, and estimation of the

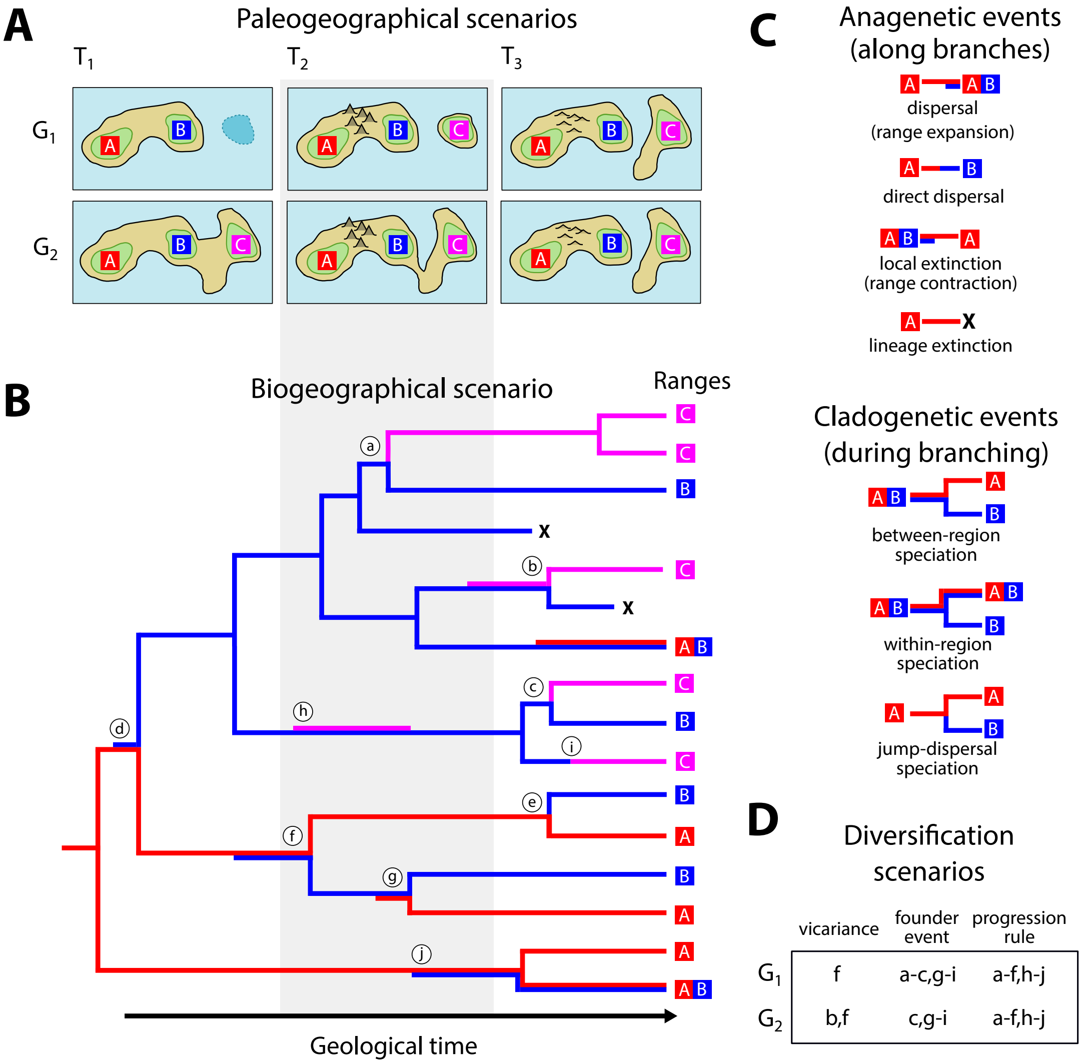


Figure 1. Anatomy of a simple phylogenetic biogeography analysis. *Panel A*: Two alternative paleogeographic scenarios (G_1, G_2) portray relationships among three discrete regions (A, red; B, blue; C, magenta) across three time periods (T_1, T_2, T_3). Under scenario G_1, regions A and B are connected during T_1; a barrier forms between A and B and an isolated region C emerges during T_2; and the barrier between A and B erodes away during T_3. Under scenario G_2, regions A, B, and C are connected during T_1; a barrier forms only between A and B during T_2; and the barrier for A and B erodes away while the land bridge between B and C submerges, forming a barrier, during T_3. *Panel B*: A realized biogeographic scenario depicting how lineages disperse, speciate, and go extinct among the three regions over time according to the paleogeographic scenarios shown in *Panel A* and through various sequences of anagenetic and cladogenetic events (*Panels C-D*). Color(s) of lineages indicate which region(s) they occupy. Terminal species are annotated with their ranges (the region(s) they occupy) or an 'x' if the species went extinct. Particular events (circled letters) are associated with different diversification scenarios (*Panel D*). *Panel C*: Anagenetic and cladogenetic events change the geographic distribution of a species relative to their ancestor and/or result in the formation of new species. Events from several methods are represented. *Panel D*: Association of different diversification scenarios with events in biogeographic scenario (circled letters in *Panel B*) under alternative paleogeographic scenarios (*Panel A*). Vicariance is favored if a new barrier causes a widespread species to split in two. Founder event speciation is favored if a species disperses over a barrier and then splits in two, or in the case of direct dispersal. Progression rule is supported when a new species disperses away from its center of origin. The Appendix contains definitions for terms.

timing of biogeographic events. We do this for several reasons. First, these goals strive to link evolutionary patterns to the tempo and mode of evolutionary processes, which is of central importance in phylogenetics and biogeography, alike. Second, since methods represent earnest attempts to reconcile the data and theory through precise definitions and rules, our focus on methods reveals much about how leading biogeographers thought about the problems of their day. Third, it allows us to celebrate the pivotal role that *Systematic Biology*, and its previous incarnation as *Systematic Zoology*, has played in advancing discussions (and airing arguments!) concerning phylogenetic biogeography. Indeed, about one of every six papers ever published in *Systematic Biology* and *Systematic Zoology* had titles and/or abstracts associated with biogeography (Appendix Figure 1).

We also note that phylogenetic biogeography has several qualities that give it a particularly dynamic history. One, phylogenetic biogeography is largely an historical, inductive, and circumstantial science. Compared to the type of experimental, deductive, and controlled findings generated by molecular biologists, it is rare to arrive at universally resolved answers in biogeography. Two, similar to paleobiology, biogeography is data-poor but implication-rich, meaning the exact results can have large consequences, but they cannot always be taken as completely certain. Three, biogeography is highly interdisciplinary and deeply ingrained within phylogenetic thinking, concerning phenomena from adaptation to speciation to extinction, so the field attracts the attention from a wide range of experts. To this last point, and in anticipation of a broad audience, we have added a visual representation of a phylogenetic analysis of biogeography (Figure 1) and a short glossary for key terms in historical biogeography to aid those unfamiliar with the jargon in our field (Appendix Glossary). For these reasons, we hope every reader will find some threads in our piece interesting.

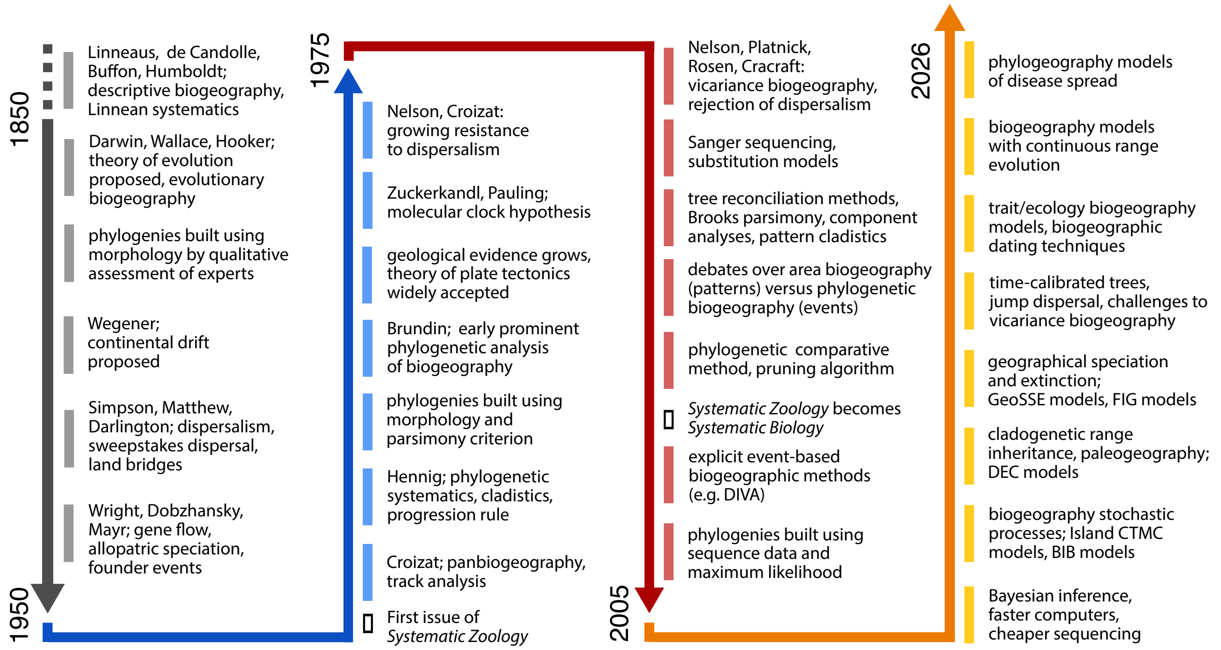


Figure 2. An evolving view of phylogenetic biogeography. Each period (arrow) corresponds to one of the four periods discussed in the main text. The ordering of themes within each period does not precisely correspond to when key ideas were introduced or popularized.

Where to begin when discussing the history of phylogenetic biogeography? To keep in theme with the anniversary of the journal, this piece mostly focuses upon the past 75 years of research. Our article provides our view on the motivations, causes, effects, conflicts, and resolutions of various research programs in phylogenetic biogeography. We have attempted to maintain balance when covering various discoveries, theories, methods, and analyses that have led to where the field is today, but we could not be exhaustive on all fronts in such a short piece. We refer readers to Hull (1990), Ebach (2015), Nelson & Platnick (1981), and de Queiroz (2014) for comprehensive overviews and varied perspectives on the history of phylogenetic biogeography. To structure the article, we bracketed the history of phylogenetic biogeography into four periods of thought: before 1950, 1950 to 1974, 1975 to 2004, and 2005 to the present (Figure 2).

Before 1950: Disjunctions, barriers, land bridges, and centers of origin

Naturalists of the 1700s and early 1800s, such as Carl Linnaeus, Alexander von Humboldt, Augustin de Candolle, and Comte de Buffon, recognized that different regions tended to be occupied by different species, regardless of environmental similarity, naturally prompting the question: Why? Since the acceptance of common descent, dating back to the Theory of Evolution by Natural Selection of Charles Darwin and Alfred Russell Wallace (Darwin & Wallace, 1858), it became clear that closely related living species tended to occupy neighboring regions, with more distantly related species living further apart. Biologists naturally sought out historical explanations for how peculiar biogeographic disjunctions – spatially discontinuous distributions separated by geographic or environmental barriers, such as the separation of Australian from South American marsupials – could have come to be. Tree-thinking and fossil evidence together helped identify where ancestral species lived, bounding when in time particular disjunctions must have arisen (Du Toit, 1937; Du Toit & Reed, 1927; Hooker, 1853), but rarely in precise or quantitative terms. Paleogeological events involving mountain-building, river capture, and epicontinental flooding were known, but the grand movements of continents, as postulated by Wegener (1912), were still not widely accepted (Kossmat, 1936).

During the first half of the 1950s, evolutionary trees remained fairly qualitative, and were generally constructed through the expert assessment of morphological evidence. Explicit phenetic and cladistic methods for reconstructing phylogenies did not yet exist, and DNA was not yet recognized as the medium for heritable genetic transmission. This is to say that biogeographers of this early period had only a fraction of the phylogenetic information we have today. Nonetheless, the population genetic views resulting from the Modern Synthesis (Huxley, 1942; Mayr & Provine, 1980) inspired new ideas for how speciation operated mechanistically within a spatial context, e.g., isolation-by distance plus time could easily produce reproductive incompatibility and lead to lineage diversification (Dobzhansky, 1937). Ernst Mayr proposed models of allopatric speciation by which new species originated following dispersal over existing barriers, or originated through subdivision caused by the formation of new geographical barriers (Mayr, 1942, 1954). Under either scenario, geographical barriers limited gene flow, thereby driving speciation. Slowly but surely, allopatric speciation became the dominant view for species formation, particularly among zoologists (Coyne & Orr, 2004). This speciation mode envisages species formation as occurring gradually and neutrally, "by chance events", rather than through selective forces, such as sexual selection or adaptation to environment (Czekanski-Moir & Rundell, 2019). From this perspective, biogeography was primarily the context for the population-level, biological aspects of speciation – one species becoming two – while taking less interest in Earth history, except as a physical template upon which allopatric, geographic speciation occurs.

There also existed an alternative view of biogeography, which focused on constructing historical explanations for regional and global distributions of species. A recurring question was: how does biogeographic history explain disjunctions between closely related lineages that are separated by seemingly impassable barriers? For example, Philip Jackson Darlington, in his synthetic view of phylogenetics and biogeography, contemplated organisms originating in a geographical area (a "center of origin") and spreading through space – colonizing new areas either passively or actively – while evolving through time. He and some members of the "New York School of Biogeography" (Nelson & Ladiges, 2001; Morrone, 2022), like William Matthew (1915), typically adopted a Northern Hemisphere-centric view of biotic dispersal, with faunal Holarctic origins and subsequent dispersal to the Southern Hemisphere. Facilitating these ideas, George Gaylord Simpson invoked the existence of geological land bridges to explain general zoogeographic patterns, such as the Great American mammal Interchange (Simpson, 1943). Simpson also convincingly argued that, while it may be unlikely for any individual species to experience "sweepstakes" dispersals over a permeable barrier, over deep timescales it becomes inevitable that at a few (not all) species complete the journey (Simpson, 1940). Exiting this period, sweepstakes dispersal and ephemeral land bridges were among the most common explanations for continental disjunctions, neither of which were very satisfying because they were each virtually impossible to falsify.

1950 to 1974: Cladistics, continental drift, and tracks

The understanding that phylogenetic relationships can provide information on the geographic distribution of organisms comes from the times of Darwin, when he proposed that organisms migrate from "centers of origin" and later evolve by natural selection to give rise to new species. Yet, it was not until Willi Hennig's cladistic method of reconstructing genealogical relationships was applied to biogeography that the field of phylogenetic biogeography was born (Cox & Moore, 2010). His seminal work, *Phylogenetic Systematics* (Hennig, 1966), laid the groundwork for the field we know today as cladistics, and revolutionized how systematic biologists approached their research (Hull, 1990). Using the principle of maximum parsimony (or Occam's razor), by which the simplest explanation is preferred in the absence of contradicting evidence, cladistics helped transition evolutionary systematics from a primarily descriptive field of natural history into (by Popper's view) a proper scientific discipline built upon falsifiable hypotheses. Hennig wrote extensively about biogeography in *Phylogenetic Systematics* and, in it, he introduced the progression rule, drawing parallels between the evolution of traits in a phylogeny and the geographic distribution of organisms: ancestral species (with more plesiomorphic characteristics) tend to remain near the center of origin for a clade, while newer species (with more apomorphic characteristics) are found in the periphery, having dispersed into newer and previously unoccupied regions.

Hennig’s "progression rule" – that the branching order of lineages in a phylogeny contains information about their geographic origins – represented a significant transformation in biogeographic practices. Early biogeographers used simple parsimony approaches, where lineages occupied one region at a time, and its location was mapped on branches or nodes to explain ancestral species location. Time-calibrated trees were extremely rare, though biogeographers such as Lars Brundin used geological information to temporally interpret parsimony-reconstructed biogeographical histories (Brundin, 1966). Thus, these methods primarily generated reconstructions that minimized the number of biogeographic movements and were considered by some as another form of "scientific dispersalism" (Nelson & Platnick, 1981). Indeed, since the second half of the 20th century, discontent grew among biogeographers regarding the ubiquity of dispersal-based explanations. If all movement over a barrier was an exception to the rule, and breaking the rule was the preferred route to explaining disjunction, what would that rule actually govern? Contemporaneous biogeographers such as Gareth Nelson (Nelson, 1978) called the New York "dispersalist school", to which Darlington, Matthew, and Simpson belonged, the "science of the improbable, the rare, the mysterious and the miraculous". While parsimony approaches for biogeography remained in a state of infancy during this early period, the next period beginning in 1975 illustrates how the versatile and exact nature of these methods soon enabled researchers to state, test, and criticize alternative biogeographic hypotheses in unprecedented (and sometimes merciless) detail.

Even before phylogenetic biogeography took root, Léon Croizat had published *Panbiogeography* (Croizat, 1958), which articulated the theory that "Earth and Life evolve together" so intimately that biogeographic patterns and paleogeographic scenarios mirror one another. To test panbiogeography theory, Croizat proposed his well-known "track analysis" method, which involves connecting species distributions on a map with lines or "tracks", and then searching for general tracks across different species as evidence of the fragmentation of an ancestral biota by vicariant events. The important work of Croizat is considered the first to vigorously challenge prevailing center-of-origin and dispersalist explanations for biogeographic patterns. Although his methodology was intensely historical, its initial methods for analysis did not account for phylogenetic relationships among species within an explicit cladistic framework; it did, however, inspire cladistic biogeographers who were also dissatisfied with dispersalism (Nelson, 1969) to develop new phylogenetic approaches to study vicariance biogeography (Nelson, 1974), as we discuss below. Although not all contemporaries accepted Croizat's views, his work has had an enduring importance, in part because it stimulated a critical re-examination of what assumptions biogeographers should and should not take for granted (Morrone, 2021).

Among biogeographers, the reality of tectonic drift remained hotly contested, as evidence remained relatively scarce and circumstantial (Du Toit, 1944; Simpson, 1943). Towards the end of this period, however, the theory of plate tectonics and the reality of drift was substantiated by new lines of chemical, geological, and physical evidence (Holmes, 1944). Accepting that the positions and adjacencies of continents had radically changed over time had immense implications for how biogeographers explained disjunctions in species distributions: rather than organisms dispersing to spread among Earth's continents, it was the continents that moved, carrying the organisms with them, "like Noah's Arks or Viking funeral ships" (McKenna, 1973). We note that, along with geological breakthroughs such as the discovery of paleomagnetism (Vine & Matthews, 1963) and the tectonic supercycle (Hess, 1962), a critical line of evidence backing plate tectonics came from biostratigraphy, where fossil sequences from currently separated landmasses aligned when considered side-by-side (e.g., the *Glossopteris* fern flora of du Toit, 1937).

Although Emile Zuckerkandl and Linus Pauling laid the foundations of molecular evolution in 1962 – highlighting that DNA not only encodes the elementary functional elements of an organism but also contains information about its evolutionary history (Zuckerkandl & Pauling, 1962) – molecular phylogenetics remained uncommon at this early stage of biogeography. The seminal work of Edwards and Cavalli-Sforza (1964) represents an early example, revealing that the hierarchical structures of genetic and geographical variation among human populations were almost perfectly congruent. For researchers working across species boundaries and on deeper timescales, phylogenies were still typically generated using morphological characters, identified and assessed using expert knowledge, and analyzed using emerging parsimony methods. The ages and relationships among many major lineages were still poorly known. The desire to locate centers of origin made ancestral area estimates of critical importance to biogeographers, since knowing where a clade originated helped interpret the historical conditions of extant lineages and their adaptations.

1975 to 2004: Vicariance, patterns and processes, molecular phylogenetics

As plate tectonics became widely accepted, continental drift joined chance-dispersal and long-gone land bridges as a plausible explanation for biogeographic disjunctions. This created new opportunities for researchers to revisit biogeographic patterns that previously defied explanation. For instance, what historical circumstances enabled species from a clade of flightless birds to be variously distributed among South America, Africa, Australia, and New Zealand (Cracraft, 1974)? Repeated rounds of overwater dispersal seemed unlikely, given what was known at the time (Mitchell et al., 2014).

This period ushered in the rise of vicariance biogeography (Nelson, 1974; Nelson & Platnick, 1981; Cracraft 1972; Cracraft 1973). In this framework, biogeographic disjunctions were explained through ancestrally widespread species that were later fragmented by the formation of a geographical or environmental barrier. Vicariance biogeographers of the time argued that the inherent stochasticity of dispersal meant it was not useful for testing most hypotheses regarding the history of biogeographic disjunctions (Croizat et al., 1974). In this sense, dispersal scenarios typically require that different species colonize new regions at random times, meaning most dispersal events must be treated as isolated cases (Nelson, 1978). If any progenitor species can disperse over any barrier at any time – such as a plant dispersing from a mainland into an island region – how can one ever falsify dispersal as an explanation for a biogeographic disjunction (Platnick, 1977)?

Vicariance, in contrast, came with two advantages over dispersal-based explanations. One, that it was falsifiable (Croizat et al., 1974; Nelson, 1974). Vicariance speciation requires that the species divides in two only after the emergence of a barrier. If the speciation occurs before (or long-after) barrier-formation, vicariance is rejected as an explanation (Nelson, 1974). Another advantage is that it consolidates evidence to test for a scenario using a "common cause". In contrast to dispersal, Croizat's axiom of Earth and Life evolving together meant that vicariance scenarios have the potential to generate a shared signal of synchronous, allopatric speciation with respect to common barriers across multiple lineages (Croizat et al. 1974).

Though Sanger sequencing was introduced in the late 1970s (Sanger et al., 1977), there were still very few densely sampled and time-calibrated phylogenetic trees. Molecular phylogenies, particularly for plant and animal clades, grew more common by the 1990s. Vicariance grew as the standard explanation for many biogeographic disjunctions, supported by geological reconstructions or evidence from the fossil record. The classical paradigm was the "Gondwanan distribution", exhibited by many plant and insect lineages found among the continents of the southern hemisphere (Brundin, 1966; Raven & Axelrod, 1972; Cracraft, 1972; Sanmartín & Ronquist, 2004).

Parsimony-based methods were extremely influential during this phase. Vicariance biogeography relied on the principle of parsimony to search for the optimal area cladogram or the most explanatory biogeographic reconstruction. Vicariance was generally assumed as the null model, whereas other processes such as extinction or dispersal were regarded as incongruities or homoplasies to be minimized during inference (Brooks & McLennan, 1991; Nelson & Platnick, 1981). Initially, three methodological frameworks dominated: component analyses (Humphries & Parenti, 1999; Platnick & Nelson, 1978), Brooks Parsimony Analysis (Brooks & McLennan, 1991), and tree reconciliation approaches (Page, 1990). Though united around the vicariance paradigm – a search for congruence between phylogenetic and distributional patterns – vicariance biogeographers disagreed on the role of geological evidence as a falsifier (Parenti & Humphries, 2004) and whether a search for general patterns was the only aim of biogeography ("area biogeography"; Ebach & Humphries, 2002), or *a posteriori* interpretation of patterns in terms of underlying events ("phylogenetic biogeography"; Van Veller et al., 2003; Wojcicki & Brooks, 2005) was a second, valid aim. Another point of discussion was the treatment of area cladograms and species distributions that departed from the vicariance paradigm, such as redundant distributions, missing areas, and widespread taxa. Tree reconciliation (Page, 1990, 1994), in which a specific cost matrix for different biogeographic events is used to map a taxon cladogram onto an area cladogram, emerged as the precursor of the next biogeographic school.

Discontent with the artificial separation between biogeographic patterns and the underlying processes triggered the birth of event-based methods (EBMs; Ronquist, 2003). If biogeographic events were ignored during the inference process, or only invoked *a posteriori* when interpreting incongruities between the general area cladogram and taxon-specific area cladograms, it became difficult to falsify hypotheses or test the significance of general patterns: the same biogeographic pattern can be explained equally well by qualitatively different historical scenarios (Sanmartín, 2012). In the event-based school, events from each biogeographic process – vicariance, duplication (speciation within a local area), dispersal, and extinction – collectively contributed to explain observed distribution patterns. The cost of each event is fixed prior to analysis using a criterion of phylogenetic conservatism: processes that maximize inherited history between ancestors and descendants (e.g., vicariance, duplication) are assigned low costs, while those that disrupt vertical inheritance—such as dispersal and extinction—are penalized with higher costs (Sanmartín, 2007). Randomization tests (i.e., randomly shuffling the tip geographic distributions or the phylogenetic relationships in the taxon cladogram) are used to find the "optimal cost" for events and to statistically discriminate between alternative biogeographic reconstructions (Ronquist, 2003). Dispersal-Vicariance-Analysis (Ronquist, 1997) represents one of the most sophisticated event-based methods of its time. However, unlike pattern-based, vicariance-based approaches, the output of DIVA is not a general area cladogram but a reconstruction of biogeographic ranges at internal nodes (range division scenarios by vicariance and duplication) and along branches (dispersal and extinction). Parsimony-based "tree-fitting (Ronquist, 2003) or "jungle" algorithms (Charleston, 1998; R. D. Page, 2003), are variants of this approach, which are also conceptually similar to various methods for host-parasite coevolutionary inference, all of which are still used today.

Thus, EBMs and cladistic pattern-based approaches (Brooks & McLennan, 1991) are alike in that they both use parsimony to infer the simplest reconstruction with the phylogenetic and biogeographic patterns, but only EBMs explicitly mapped events onto the phylogeny to generate biogeographic histories consistent with observed patterns (Sanmartín, 2012). Despite being grounded in a more explicitly mechanistic framework, EBMs have been criticized by some vicariant biogeographers, under the view that EBMs are overly reliant upon dispersalist or center-of-origin histories (Crisp & Cook, 2005; Ebach et al., 2003) or for assuming a single area history, thereby ignoring species contingencies (Wojcicki & Brooks, 2005). However, these critiques largely reflect differences in interpretation rather than inherent limitations of the framework. EBMs helped shift biogeography away from the practice of inferring ancestral regions by aligning phylogenetic and biogeographic patterns, and towards investigating the phylogenetic placement of events needed to produce a biogeographic pattern (Sanmartín, 2022).

However, still missing among parsimony methods was an explicit temporal dimension. Taxon sampling and time-calibration approaches improved toward the end of this period, creating compelling evidence for long-distance dispersal in the history of many lineages. For instance, Raven and Axelrod, through consultation of the fossil record, were quite aware that the relatively young ages and tropical disjunctions of many angiosperm lineages were incompatible with continental vicariance explanations (Raven & Axelrod, 1974), as Pangaea and Gondwana were largely broken up by the Cenozoic, with only relatively localized geological changes occurring thereafter (e.g., the formation of the Isthmus of Panama, opening of Drake's passage, etc.). In addition, because parsimony methods generally do not incorporate time into the analysis, they were prone to conflate biogeographic congruence (true positive, correspondence in spatial and temporal patterns) with pseudocongruence (false positive, identical spatial patterns with different temporal origins) and even pseudoincongruence (false negative), all of which require temporal information to distinguish (Donoghue & Moore, 2003).

2005-2025: Timing, probability, and models

The smaller datasets of the previous era meant biogeographers were often limited to considering deep "backbone" disjunctions and movements in their analyses. By the early 2000s, plummeting sequencing costs granted biogeographers access to phylogenies of unprecedented quality, with denser taxon sampling and greater topological resolution. Public databases, such as GenBank (Bilofsky et al., 1986) for genetic sequences, GBIF (*Global Biodiversity Information Facility*, 2001) for geospatial occurrences, and the Paleobiology Database (Alroy et al., 2001) for fossil occurrences, grew steadily. These data, combined with advances in molecular dating methodologies (Thorne & Kishino, 2002), enabled biologists to rapidly produce time-calibrated phylogenies and species range matrices for larger, densely sampled clades. With improved phylogenetic and biogeographic resolution, researchers became increasingly capable of estimating the tempo, mode, and timing of dispersal and speciation in precise, quantitative terms (Renner, 2005; but see Heads 2005 & 2012). In addition to helping distinguish biogeographic pseudocongruence from true shared history (Donoghue & Moore, 2003), time plays a key role in understanding propensities for biogeographic change. When branch lengths in a phylogeny reflect evolutionary change, they can also inform us about the likelihood of biogeographic events, and what should be considered "rare". Recalling Simpson’s view on the inevitability of "sweepstakes dispersal", longer branches necessitate more time and thus a greater chance for rare events of dispersal, extinction, or for a lineage to be influenced by geological processes.

Much of this progress was made possible by newer process-based probabilistic models that built on the successes of the likelihood revolution in molecular phylogenetics (Felsenstein, 1981). These models use stochastic processes to approximate the behavior of biogeographical processes – such as dispersal, extinction, and speciation – in the production of biogeographical events, both in terms of timing and outcome. Using maximum likelihood or Bayesian inference approaches to fit these models to phylogenetic and biogeographic patterns allowed researchers to compare rate estimates, reconstruct ancestral ranges, and statistically test competing biogeographic hypotheses. Lastly, these models were also used to simulate biogeographic patterns, which was crucial for assessing model realism, validating method accuracy, and exploring hypothetical biogeographic scenarios. The randomness of probabilistic models helped capture the uncertainty inherent to historical inference problems: rather than reconstruct the single most-parsimonious history, models instead reconstructed sets of (probabilistically) plausible histories (Ree & Sanmartín, 2009).

The first biogeographic models adapted continuous-time Markov chains (CTMCs), commonly used to model discrete traits (Pagel, 1999), for application in island biogeography. In these Island CTMC models, lineages move between pairs of regions through direct dispersal (A→B, B→A) during anagenesis, with species occupying one region at a time. Nepokroeff et al. (2003) was among the first to apply maximum likelihood methods to study biogeography, applying the approach to Hawaiian *Psychotria*, while remarking on the need to account for estimation uncertainty when making biogeographic inferences, the importance of branch lengths (time) for understanding the probability of biogeographic change, and the potential for adding realism to models to improve estimates. Biogeographers soon realized that the simplicity of CTMCs makes them highly adaptable and scalable for modeling complex biogeographic scenarios, simply by modifying the structure of the underlying rate matrix.

An extension of the Island CTMC approach is the Bayesian Island Biogeography (BIB) model (Sanmartín et al. 2008), which was developed to infer general patterns of geographic movement across multiple clades radiating among a shared set of islands. This is achieved by introducing hyperparameters into the Island CTMC rate matrix that govern range evolution across clades – such as between-island dispersal rates and island carrying capacities – while clade-specific traits (e.g., tree topology, branch lengths, molecular substitution rates, age of origin, and migration rates) are modeled through individual parameters that are jointly marginalized over as part of Bayesian inference.

As a Bayesian approach, BIB explicitly accounts for two major sources of uncertainty often disregarded by parsimony and maximum likelihood-based approaches: phylogenetic and parameter uncertainty. Early efforts to incorporate phylogenetic uncertainty applied parsimony-based biogeographic methods to samples from the Bayesian posterior distribution of trees (Huelsenbeck & Imennov, 2002; Nylander et al., 2008). In contrast, BIB jointly estimates the phylogeny and parameters governing molecular and biogeographic evolution (e.g., substitution rates, migration rates) directly from sequence alignments and species location data, thereby integrating phylogenetic uncertainty throughout the inference process. Parameter uncertainty is likewise accommodated, as all parameters are inherently estimated as posterior densities rather than point estimates, capturing the notion that we do not (and should not) claim to know the exact rate of dispersal between regions (Sanmartín et al. 2008).

Despite all their advantages, models within the Island CTMC family had several shortcomings for modeling species-level biogeography. First, the ranges of widespread species cannot be represented using a single-region encoding. Second, these models ignored the important influence of speciation in shaping the range of newly formed species. Third, they assumed dispersal rates between regions were constant over time and ignored periods when particular regions (e.g., volcanic islands) did not exist (Sanmartin, 2022). Amending these issues, the Dispersal-Extinction-Cladogenesis (DEC) model (Ree et al., 2005; Ree & Smith, 2008) can be viewed as a probabilistic treatment of DIVA that fundamentally differed from Island CTMC models. DEC allowed species to occupy multiple regions simultaneously, daughter lineages to asymmetrically inherit species ranges following cladogenesis, and paleogeographical information (e.g. paleocontinental adjacency) to shape dispersal rates as a function of time. DEC also distinguished itself from DIVA, which generally assumed vicariance was responsible for lineage divergence, by instead treating cladogenesis as the random outcome of speciation, either within or between regions. Anagenetic range evolution along phylogenetic branches is modeled in DEC as a CTMC with two event types. Unlike the simple region-switching movements of Island CTMCs (A→B, B→A), dispersal is modeled as range expansion (A→AB), while geographic extinction is equated to range contraction (AB→B). Consequently, a species in region A must first expand into region B and then contract from region A in order to "shift" its entire range from A to B under the DEC model.

Among DEC’s many features, the cladogenetic component was particularly captivating. In 2005, relatively few phylogenetic models allowed asymmetric trait inheritance during time of speciation (but see Bokma, 2002). DEC supported two cladogenetic scenarios defined previously by DIVA: Scenario 1 involved an ancestor in one region yielding two new species in that region (A→A,A) and Scenario 2 involved a widespread species being split into two regions (AB→A,B). DEC cladogenesis differed from DIVA by introducing Scenario 3, which allows the partial inheritance of widespread ranges (AB→AB,A), and by disallowing vicariant speciation patterns (ABCD→AB,CD) modeled by DIVA, on the basis that DEC was not equipped to adequately model vicariance. Even though Ree et al. (2005) was careful to define cladogenetic events purely in terms of state patterns, practicing biogeographers came to interpret the probability for certain cladogenetic inheritance patterns as evidence to support or reject particular speciation modes. For example, the Scenario 2 pattern involving a split in ancestral ranges is often equated with vicariance, even when no causal paleogeographic event is referenced.

During this era, biogeographers still wished to test whether biogeographic disjunctions were predominantly caused by vicariant or founder event speciation. This motivated Matzke (2014) to introduce the DEC+j framework, an extension of DEC that includes jump-dispersal: a new cladogenetic event, distinct from the anagenetic dispersal process, that allows new species to acquire a region outside the ancestral range during speciation. Whereas previous studies used statistical model choice to measure support for alternative anagenetic dispersal models (Ree & Smith, 2008; Sanmartín et al. 2008), Matzke (2014) instead used model selection to measure support for alternative cladogenetic model variants. Particularly in cases where each species tends to occupy a single region (e.g. island systems), biogeographers have found that DEC+j is frequently selected over standard DEC models that lack jump-dispersal cladogenesis. While biogeographers generally expect that dispersal over a barrier will sever gene flow and eventually lead to speciation, whether DEC+j is suitable for modeling this complex phenomenon, especially over deep time scales, has been both critiqued (Ree & Sanmartín, 2018; Sanmartín, 2022) and defended (Matzke, 2022) on biological and statistical grounds. Exactly how to model jump dispersal – whether as a compound event of dispersal-with-speciation, akin to an instantaneous founder-speciation event, or as a sequence of events, caused by dispersal over a barrier followed by allopatric speciation – remains an open question warranting further discussion and study (Ree & Sanmartín, 2018; Matzke, 2022; Landis et al. 2022).

Central to questions of cladogenesis is the fact that Island CTMC and DEC models are trait variation processes, not branching processes that generate trees: they take the phylogenetic relationships and branching times as given, and then "map" range evolution histories onto that tree. This has several consequences. First, cladogenetic events in DEC (and DEC+J) are modeled without a temporal context (reminiscent of parsimony approaches) which can lead to undesirable properties. For example, because DEC models do not allow species range to influence speciation rate, a widespread species whose range crosses a barrier does not experience an elevated risk to undergo allopatric speciation, meaning that species can maintain unrealistically fragmented ("spatially disjunct") ranges indefinitely. Second, DEC models will underrepresent cladogenetic range evolution resulting from "hidden" speciation events that left no observed descendants due to e.g. extinction or undersampling (Cornuault & Sanmartín, 2022), making them sensitive to incomplete or uneven taxon sampling. Third, DEC does not allow species to truly go extinct, but rather approximates extinction by placing imperiled species into an inescapable absorbing state called the "null range" (a range with no regions; Massana et al., 2015). In effect, this means that DEC paradoxically expects some extant species to possess (extinct) null ranges, leading the model to underestimate extinction rates and overestimate the importance of cladogenetic events when explaining why species have narrow ranges.

Concerned with these issues, Goldberg and collaborators modeled how dispersal, speciation, and extinction events act upon species ranges to directly produce phylogenetic and biogeographic patterns (Goldberg et al., 2011). Their geographic-state speciation-extinction (GeoSSE) model translated the general logic and event-set of DEC into the state-dependent speciation-extinction framework of Maddison et al. (2007). Under GeoSSE, species evolve through dispersal (range expansion), local extinction (range contraction), within-region speciation, and between-region speciation, where all rates depend on the species ranges before and after the event. Consequently, species ranges influence biogeographic rates, and biogeographic rates shape species ranges. GeoSSE yielded several improvements over DEC, such as the ability to test whether speciation or extinction rates varied among regions or ranges, and a mathematically explicit way to account for "hidden" speciation events. In contrast to DEC, which could not model relationships between local extinction and lineage-level extinction, GeoSSE linked the two events, defining lineage-level extinction as local extinction in the last region of the species range.

Although GeoSSE allowed species ranges to instigate speciation, it provided no inherent structure to predict how variation among regional features (e.g. sizes, distances, elevations) should shape biogeographic outcomes. Taking allopatric speciation as an example, species that disperse over barriers should quickly divide in two due to lack of gene flow. Those same barriers should also limit how often species move among regions. Indeed, many regional features (e.g. size, elevation, temperature) are predicted to shape rates of speciation, extinction, and dispersal, with the effects of those features on rates differing from clade to clade. Feature-Informed GeoSSE (FIG) approaches model the sign and magnitude of relationships between rates and regional features (Landis et al., 2022; Swiston & Landis, 2025) as part of the tree and range evolution process. For instance, whereas DEC models assume that all widespread ranges undergo allopatric speciation (by splitting the ancestral range) with equal probability, FIG uses the distribution of barriers among its regions to inform the rate at which widespread species with unstable ranges split into daughter lineages with stable ranges (Landis et al., 2022).

How to model biogeographic rate variation and, therefore, which spatial factors govern those event rates, has always been of acute interest and importance to biogeographers. However, parsimony-based biogeographic approaches required users to assign costs to different event types, which becomes difficult to do in a principled manner when there are large numbers of unknown factors. For example, what cost should be assigned to speciation occurring in a flighted species that is widespread throughout a tropical mountain range relative to the cost of extinction of a wet-adapted species in a desert? Probabilistic models, on the other hand, estimate what these costs should be from the dataset in question and use the laws of probability to assign those costs in a coherent way to alternative biogeographic scenarios. As such, developers of methods have proposed dozens of new phylogenetic approaches to model how biogeographic rates are informed by phenotypic traits (Sukumaran et al., 2016, Klaus & Matzke, 2020), unobservable "hidden" traits (Caetano et al., 2018), paleogeography (Buerki et al., 2011), paleohabitats (Landis et al., 2021a; Quintero et al., 2023), competition (Quintero & Landis, 2020), and the geometry of regional features (Cardillo et al., 2017). Models exploring how various geographical and ecological factors shape biogeographic rates are too various to summarize here. However, they share a common inference strategy and objective: to incorporate information beyond the standard phylogenetic and biogeographic inputs to infer how those extrinsic factors shape biogeographic rates.

Not surprisingly, any results from phylogenetic methods for biogeography depend critically upon the evidence at hand. However, what is "known" versus "missing" varies wildly among biogeographical systems (e.g., temperate mammals versus deep-sea microbes), due to gaps in our knowledge of taxonomy, evolutionary relationships, and geographic distributions

(Hortal et al. 2015). For instance, incomplete taxon sampling, particularly when biased to exclude certain regions (e.g. remote islands), is prone to underrepresent the true extent of range evolution in a clade. Similarly, incorrect phylogenetic relationships can create spurious patterns of biogeographic disjunction, leading to overestimated rates of species movement. Taxonomies based on "oversplitting" or "overlumping" lineages artifactually alter patterns of species richness and endemicity, which would also distort biogeographic inferences. False presences and/or false absences in range maps and georeferenced species occurrence databases could also misinform inferences. Despite its importance, strategies to handle uncertainties, biases, and errors in the input data remain relatively underexplored.

As noted earlier, this most recent period witnessed an explosion of biogeographic analyses, sustained by new modeling techniques and growing numbers of fossil-dated molecular phylogenies. Although early molecular systematists voraciously incorporated temporal information from fossil taxa to date their phylogenies, spatial information about ancestral lineages often went unused. Paleobiogeographers warned that ignoring fossil evidence in biogeographic studies is dangerous, as past episodes of dispersal and extinction may cause surviving, descendant species to inhabit regions completely apart from their ancestral origins (Lieberman, 2002). Since then, researchers have introduced various techniques to incorporate fossil biogeography into reconstructions, with early examples including parsimony methods (Swenson et al., 2001), fossil node "priors" for ancestral ranges (Meseguer et al., 2015), fixing fossils tips within a backbone phylogeny (Mao et al., 2012; Wood et al., 2013), or even treating fossil taxa as part of the phylogenetic inference procedure (Landis et al., 2021b). Including fossil taxa is easier said than done, due to the notorious incompleteness of the fossil record

(Lieberman, 2002). The incompleteness itself varies in intensity across different times, places, and clades, due to differences in rates of fossil preservation, extraction, and description (Benson et al. 2021). While there exist non-phylogenetic methods of historical biogeography that account for spatiotemporal variation in fossil richness (Silvestro et al., 2016), we lack explicitly phylogenetic models that do so. When the fossil record delivers, such as for the case of primates (Wisniewski et al. 2022), researchers have convincingly shown that ancestral ranges reconstructed solely from extant species are prone to underrepresent dispersal and misestimate the place of origin. Even taking a single fossil into consideration has radically altered how we understand the historical biogeography of groups such as hummingbirds (Mayr 2004) and daisies (Barreda et al. 2015). Conversely, biogeographic analyses of entirely extinct groups, such as non-avian dinosaurs, have altered how we understand their historical movements and diversification (Griffin et al. 2022). This is to say that, among biogeographers and paleobiologists, there is a firm consensus that fossils should be included in phylogenetic ancestral range estimates whenever possible, though more work in this area is needed.

For clades lacking strong fossil records (e.g. many plant and insect groups), systematists have long combined biogeographic scenarios with paleogeographic evidence to date divergence times for disjunct lineages (Brundin, 1966). For example, one might constrain the maximum age of a clade to be younger than the volcanic island it inhabits. Using biogeographic scenarios to date nodes *a priori* is not without complications, as it requires the researcher to first observe cladogenesis in the tree and then presuppose the timing and nature of the historical scenario producing the split (Baldwin & Sanderson, 1998). Moreover, any phylogeny time-calibrated in such a manner is tainted and cannot be safely re-analyzed a second time to study historical biogeography of the clade (Kodandaramaiah, 2011; Renner, 2005). Recent work has shown that these problems of uncertainty and circularity can be avoided by jointly modeling a unified phylogenetic, biogeographic, and paleogeographic history (Landis, 2017; Landis et al. 2018). Under such a framework, all modeled historical scenarios are considered, where those of high likelihood (e.g. a single colonization event into a young island followed by a recent radiation) are favored over less parsimonious scenarios (e.g. an ancient radiation followed by ten independent island-colonization events). As such, ancestral range and divergence time estimates are inferences from, rather than assertions upon, the dataset and model.

Coming full circle, founder event or jump-dispersal speciation, deemed by vicariance biogeographers as untestable (Croizat et al., 1974), returned to prominence during the period thanks in large part to the proliferation of time-calibrated trees, the advancement of biogeographic methods that explicitly account for geological time, and the adoption of statistical hypothesis testing methods (de Queiroz, 2014; Crisp et al., 2011). Denser phylogenetic sampling and methods that are temporally explicit revealed that many biogeographic disjunctions were created, not by ancient vicariance events, but recent dispersal events over barriers (Cook & Crisp, 2005). As a result, vicariance drifted from being the favored hypothesis for explaining most disjunctions, to one that is in an even competition with many newer, dispersal-based hypotheses. On the other hand, paleobiogeographers who incorporate fossil taxa and Earth history data into their analyses routinely recover support for ancestral vicariance events (Griffin et al., 2022). Confidently discriminating between these competing historical scenarios ultimately depends on accurate estimates for the timing of paleogeographic barrier formation, biogeographic dispersal, and phylogenetic lineage divergence. The assembly of newer, integrative datasets with comprehensive sampling of living and extinct species, richer paleogeographical information, and the invention of more nuanced statistical tests, may help identify which disjunctions were caused by the appearance of a barrier (vicariance) versus the dispersal over a barrier (founder event).

It is also important to stress that cladistic and event-based biogeography tested hypotheses of vicariance by searching for shared patterns across multiple clades (Marshall & Liebherr, 2000; Sanmartín et al., 2001). The advent of more sophisticated, parameter-rich biogeographic models from the DEC and GeoSSE families have returned focus back to single-clade approaches, reminiscent of Brundin (1966). As we have seen, BIB can be used to detect shared patterns while accounting for lineage-specific components of biogeographic history in a set of co-distributed taxa, and has been extended to the study of "continental island" patterns (Sanmartín et al. 2010). Other diversification models within the DAISIE framework were developed to test general theories of island biogeography (e.g. diversity-dependent diversification) by estimating parameter across multiple island clades (Valente et al. 2015, 2020). However, the multiclade analytical approaches used in BIB and DAISIE have not yet been extended to explicitly model vicariance in multiclade systems. Rather than use true multiclade models, biogeographers instead combine results from single-clade models, either across many smaller clades (Ye et al. 2023; Ding et al. 2020; Merckx et al., 2015) or within large individual clades (Wootton et al. 2025), to identify "common causes" underlying biogeographical movements and disjunctions. This suggests that vicariance continues to interest modern biogeographers, as it had for earlier cladistic-vicariance biogeographers of the last period, and that there is an unsatisfied appetite for new statistical methods in this area.

All phylogenetic methods of biogeography discussed so far assume, if not hope, that species ranges can be realistically represented as discrete characters, e.g. as sets of occupied regions. This simplicity is convenient for defining models, generating datasets, and testing hypotheses. Keeping the number of regions small is crucial, as most computational methods have struggled to analyze systems with unconstrained range sizes for more than (say) 10 regions (Landis et al. 2013). For these reasons, the practice of demarcating regions is rather subjective and system-specific. In addition, the use of large and poorly-defined regions would likely fail to capture the fine-scale nature of geographic variation. These shortcomings of discrete methods, combined with the explosion of high-resolution geospatial datasets, fueled the rise of new biogeographic models that represent species locations in continuous space. These approaches typically use Brownian diffusion as the driver of movement (Lemmon & Lemmon, 2008) and represent species locations with a geographical coordinate or a point (e.g. a sample or the centroid from the range). Continuous models have been invented to allow diffusion speed to vary among lineages (Lemey et al., 2010), to integrate species range data (Nylinder et al., 2014; Quintero et al., 2015), to correctly represent species movements on spherical surfaces (Louca, 2021), to allow for the subdivision of ranges through the stochastic appearance of barriers (Albert et al., 2017), and to account for paleogeography when reconstructing movement (Arias, 2024). This leaves an open question, do discrete representations of biogeography (i.e. using "areas" or "regions") serve a purpose or are they a holdover from parsimony-era analysis? Despite the advantages of explicitly representing species ranges in continuous space, treating all directions of movement as equal—as in the Brownian (random-drift) model—can obscure general dispersal trends. In particular, rare between-area dispersal events may be masked by the much more frequent within-area movements under a diffusion model (Ronquist & Sanmartín, 2011). Ideally, future biogeographic models that represent species locations in continuous space can be adapted to account for phenomena currently captured by existing, discrete-based models, such as location-dependent speciation and extinction or cladogenetic range inheritance.

As we conclude this section and reflect on recent times, phylogenetic biogeography has proven relevant to evolutionary phenomena that play out on human timescales: namely epidemiological outbreaks and the spread of cancers. The CTMC models, originally used to study island biogeography, have been adapted to model pathogen phylogeography and epidemiology, with early work in this area pioneered by Lemey et al. (2009). In their model, viruses spread among populations (e.g. of countries) during an outbreak, while using Bayesian techniques to favor pathogen spread among the smallest set of short-distance routes. Other innovations include modeling migration rates to vary across calendar years (Bielejec et al., 2014), hierarchically modeling shared transmission tendencies across independent outbreaks (Cybis et al., 2013), and using generalized linear models to use external predictors (e.g. distances) to inform migration rates (Faria et al., 2013). Using techniques such as these, epidemiologists now routinely analyze pathogen sequences to track the origin and spread of infectious disease, ranging from recent outbreaks, such as those for SARS-CoV-2 (Bedford et al., 2020; Nadeau et al., 2021) or seasonal influenza (Bedford et al., 2015), to historical scenarios, such as the introduction of HIV into the US (Worobey et al., 2016). In terms of other evolutionary diseases, biogeographic approaches have also been used to model the spatial diversification of cancerous cells within a tumor (Lewinsohn et al., 2023) and the spread of cancer among tissues during tumorigenesis (Alves et al., 2019; Chroni et al., 2021). Because both epidemics and cancer-spread can be completely observed (sequenced) on human timescales, and because these may be accompanied by a wealth of patient or host data (e.g. risk factors), they are ideal for studying many properties of biogeographic systems. Knowing how the past shaped the present should help predict the future, so understanding the temporal, spatial, and phylogenetic context of evolutionary disease holds immense promise to improve human quality of life.

Future growth

Due to its versatility and exactness, our view is that the current process-based modeling paradigm holds the most promise to advance phylogenetic biogeography. However, it is still far from reaching its full potential, largely because existing models are primarily informed by phylogenetic patterns and geographical occurrences. What else might be missing?

Models, of course, represent a field's humble attempts to formally describe nature as we observe it. Most phenomena we wish to model ultimately derive from the experience and expertise of organismal biologists, who ideally represent the study of a diverse range of distinct lineages and geographical settings. The truth is not all clades and places have been adequately represented by our current approaches. Consider, for instance, that many of the key figures (including Darlington, Simpson, Brundin, Rosen, Platnick, and Nelson) involved in the early debates published in *Systematic Zoology* were, of no surprise, zoologists. Could their taxonomic focus have led them to place geographical barriers above, say, environmental filters and biotic interactions when formulating their ideas behind dispersalism and vicariance biogeography? Imagine if instead botanists, following in the ecological traditions of de Candolle and Humboldt, laid the foundation for phylogenetic biogeography that emerged throughout the 1960s and into the 1980s. Perhaps our theoretical frameworks would instead focus on range evolution in response to climate, soil type, wind and sea current, and interspecific interactions with pollinators and seed-dispersers. Modern botanists of the molecular era might have first extended the theory to incorporate the reciprocal effects between hybridization and polyploidy with biogeography, and so on. Going further, how would phylogenetic biogeography appear today, had its central ideas come primarily from mycologists, from microbiologists, or from virologists?

While there are no correct answers to this type of thought experiment, it helps remind us there are many "blind spots" persisting in our field. Moving forward, we should strive to be more creative regarding what additional lines of evidence – including from population genetics, paleobiology, ecology, and morphology – we consider when constructing or testing our biogeographic hypotheses.

Phylogenetic models of biogeography would benefit immensely by incorporating ideas from mechanistic, population genetic models used to study phylogeography (Avise, 2000). Sequence data availability was a limiting factor 20 years ago, but clade-wide population-level sequencing is now affordable for many researchers. In particular, incorporating phylogeographic techniques can generate new information about the nature of organismal movement versus gene flow (Levin, 1981), patterns of within- and between-species population structure to infer allopatric speciation (Choi et al., 2021), the use of historical population bottlenecks to date and locate past dispersal and/or extirpation events (Vincek et al., 1997), how alleles are inherited under alternative speciation scenarios (Avise, 2000; Knowles & Maddison, 2002), and how shifts in allele frequencies indicate adaptation to novel environmental conditions (Coop et al., 2009).

Integrating this microevolutionary perspective into future model design will be important for advancing interspecific approaches to phylogenetic biogeography.

That said, a microevolutionary perspective that focuses too intently on shallow timescales, at the exclusion of deeper timescales, might miss the big picture. Most studies of phylogenetic biogeography are already neontological, relying solely upon the distributions and relationships among extant taxa to make historical inferences, even though extant taxa may differ from where or how their shared ancestors lived (Faurby et al., 2024; Wisniewski et al., 2022). Phylogenetic biogeographers interested in realistic historical inferences should make the effort to incorporate the fossil record and Earth history data whenever possible. Developing models that explicitly account for the geographical and stratigraphic ranges of ancestral species (e.g. as chronospecies; Hopkins et al., 2018), the unevenness in fossil sampling effort (Silvestro et al., 2016), the uncertain phylogenetic placement of microfossils (Barreda et al., 2015), fossil-based evidence of extirpation (Jones et al., 2009), and higher-resolution paleogeographical and paleoenvironmental data (Scotese, 2021) will be crucial for understanding how biogeographic processes operate at the deepest timescales.

Just as integrating vertically with respect to timescales will be key, so will it be to integrate horizontally across lineages in the ecological dimension. Most event-based models in phylogenetic biogeography assume that species disperse, originate, and go extinct more-or-less randomly, not as the consequence of species coexistence or environmental change (Vellend, 2010). This is partly reflected in the fact that models of phylogenetic biogeography have primarily focused on "habitation" (e.g. geographical range) while largely ignoring "station" (e.g. environmental conditions for survival). How changing spatioenvironmental conditions influence species persistence, dispersal, extinction, and adaptation, particularly over deep timescales, is of central importance, but vastly underexplored from the modeling perspective (Donoghue & Edwards, 2014). Taking this further, biotic interactions resulting from competition for limited resources (MacArthur & Wilson, 1963), mutualistic plant-pollinator associations (Hembry et al., 2013), and parasitic dependencies upon hosts (Hoberg & Brooks, 2008) all shape biogeographic processes and patterns, both in terms of habitation and station. Even though modeling how historical ecological interactions form and dissolve in space over time within a phylogenetic context is not easy, advances in this area will be important for the field.

Variation in morphological adaptations also have a profound impact on where different species disperse, survive, and diversify (Sukumaran & Knowles, 2018). The presence or absence of categorical traits, such as powered flight among birds, undoubtedly impacts many dispersal patterns (Garcia-R & Matzke, 2022). Similar but more complicated factors, such as variation in airborne seed buoyancy in relation to atmospheric circulation (Vasconcelos et al. 2021), must also impact dispersal, but remain difficult to model explicitly. There is no shortage of other fascinating traits of biogeographical relevance – e.g., among angiosperms: wood anatomy, leaf succulence, seed dormancy, fruit color, deciduousness, the ability to self-fertilize, etc. – yet the few current approaches we have that account for morphology inspect only a single trait, rather than consider the integrated effects of all traits at the organismal level. The ability to model how entire suites of evolving organismal traits collectively shape, and are shaped by, range evolution will be an important step forward.

That being said, there is no simple or singular recipe to elevate phylogenetic biogeographic research to its next, logical level. Most empirical research focuses on the analysis of species-level phylogenetic trees with geographical ranges. Combining standard phylogenetic biogeography analyses with comparative analyses of population-level genetic variation, functional traits, climatic niches, ecological interactions, and/or fossil data will continue to be instrumental for strengthening evidence surrounding basic biogeographic principles, and for identifying weak points, contradictions, and unsolved puzzles that deserve critical attention. It is this type of careful empirical work that inspires the design of new modeling tools used by the next generation of biogeographers, creating a positive feedback loop for future discoveries.

Integrating microevolutionary scales into macroecological questions, together with the incorporation of additional independent sources of evidence as noted above, has resulted in new and complex models that are both increasingly realistic and parameter-rich (Skeels et al., 2019; Overcast et al. 2021; Hagen et al., 2021). This added complexity translates not only into substantial computational demands, but also, more importantly, risks the construction of models with intractable likelihoods and/or non-identifiable parameters. In this context, emerging applications of supervised machine learning (Voznica et al., 2022; Lambert et al. 2024, Qin et al., 2026; de la Peña et al., 2026) and generalizable modeling tools (Landis & Thompson, 2025) that are increasingly being used for complex biogeographic inference problems, such as the spread of pathogens in space (Thompson et al. 2025), the effects of local species richness on diversity-dependent biogeography (Soewongsono & Landis, 2026), and the reconstruction of ancestral locations (Nagel & Landis, 2026), and may expand the range of biogeographic phenomena and scenarios we explore in the near future.

For most biogeographers, the very aspects of phylogenetic biogeography that complicate its study are also what make it such an alluring subject. The grand questions that engrossed (and vexed!) early dispersalist and vicariance biogeographers of the last century brought us to where we are today. Continuing in this tradition, our work as phylogenetic biogeographers is to understand what processes generated the spatial history of all life in all possible places over all possible timescales. Fuller integration of microevolutionary, paleobiological, ecological, morphological, and other perspectives into our empirical studies and our models will undoubtedly bring us closer to a fully explanatory theory of biogeography. Despite how promising the next generation of datasets, models, and analyses might be, we should not hope to answer our questions in biogeography through individual efforts. To advance phylogenetic biogeography in particular, and biogeography as a whole, we must strive to build truly multidisciplinary collaborations, involving organismal experts, geneticists, ecologists, climatologists, geologists, computational biologists, and statisticians, to make progress in this captivating area of natural history research.

Acknowledgements

We are extremely grateful to the Society of Systematic Biologists' Legacy Committee for the opportunity to share this article as part of the 75th volume of *Systematic Biology*. We are also indebted to two anonymous reviewers, James Albert, Bob Thompson, and members of the Landis Lab for comments that helped improve the quality of the manuscript.

Conflicts of Interest

None declared.

Funding

MJL was supported by funds from the National Science Foundation (DEB 2040347) and the National Institutes of Health (NIH R01-TW012704). IS received funding from project PID2023-153023NB-I00 by the MCIN/AEI/10.13039/501100011033/ and European Union NextGenerationEU/PRTR.

Author Contributions

MJL and IS conceptualized and wrote the manuscript. All authors reviewed and approved of the content in the manuscript.

Appendix

Glossary

Brief definitions of terms used in the associated manuscript, and in phylogenetic biogeography more broadly, are provided below. See Figure 1 in main text for visualizations.

- *Range*: The basic unit of biogeographical analysis. The range represents the location, size, shape, and extent of the geographical distribution of a taxon (e.g., a population, a species, a clade, etc.). Ranges may be encoded as sets of occupied regions, individual localities, centroids, rasters, or polygons. The choice of range encoding is influenced by data availability and desired precision (e.g., expert-based range maps versus computer-determined boundaries), methodological requirements (e.g., input as region-sets or continuous ranges), and the biological implications of the encoding (e.g., based on species ranges versus geological features).

- *Barrier*: A geographical (e.g., an ocean) or environmental feature (e.g., uninhabitable land lying outside a given species' climatic tolerances) in the physical template of Earth that impedes regular movement into new regions and/or interrupts the gene flow between two populations or species. Barriers can be permanent or temporary, they can be semi-permeable or impassable, and they can change in location and quality over time. Not all species interact with a barrier in the same way. In fact, a barrier for one species might be a dispersal route for another (e.g. a river).
- *Region*: A defined area of the Earth that is delimited by geographical or environmental features (e.g. barriers), patterns of endemism (e.g. uniqueness in species composition), or other criteria. Region definitions often depend on the scale of the analysis and the hypotheses being tested. For example, a study of dispersal among Caribbean islands might define each island as one region, whereas a global study of continental radiations might lump all Caribbean islands into one region.
- *Endemic and widespread species*: Endemic species live in only one region whereas widespread species live in many regions. Endemic species are often assumed to be more specialized to local condition and at higher risk of extinction, whereas widespread species are often considered to be more generalist and at lower risk of extinction.
- *Disjunction:* The separation of one or more species' ranges into geographically distinct regions by a barrier. Biogeographically disjunct populations within a species are expected to share lower levels of gene flow and be at higher risk of experiencing

allopatric speciation. Biogeographically disjunct species refers to two or more species (often sister lineages) being isolated into different regions. For example, a continental species that colonizes an island will afterwards have disjunct populations. If the disjunction leads to allopatric speciation, the two species will represent a biogeographic disjunction, while the populations within each species will no longer carry the disjunction that led to speciation.

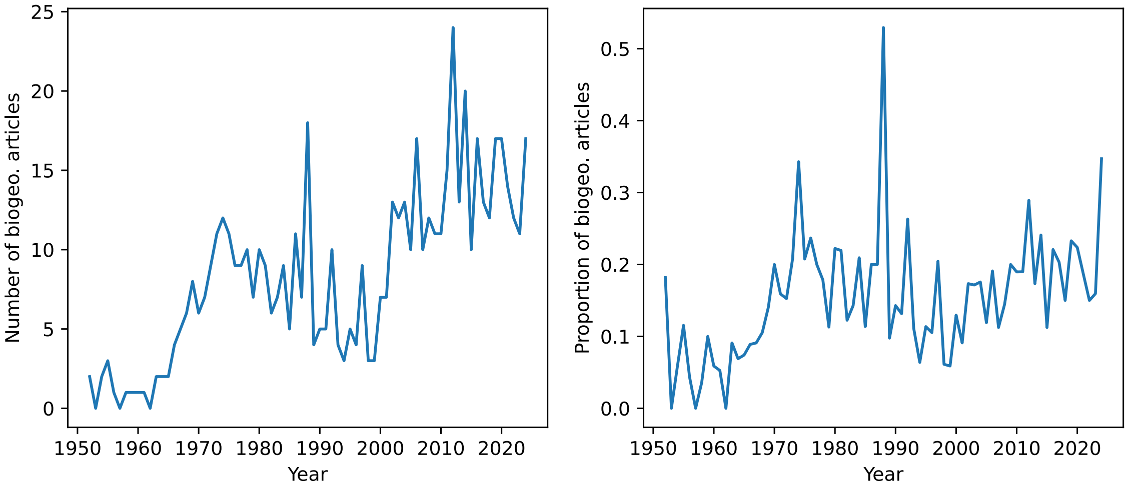
- *Dispersal*: A biogeographic event in which a species establishes a new population in a location outside of its ancestral range. Dispersal expands a species' range if the newly established population(s) maintain(s) enough gene flow with the ancestral populations to prevent lineage divergence, causing a species to become more widespread; this process is sometimes referred to as "range expansion" or "dispersion". Alternatively, dispersal over a barrier to gene flow may result in founder event or jump-dispersal speciation (see also *Vicariance and founder event speciation*). Direct dispersal, where a lineage switches its occupancy from one region to another, has also been used to represent founder-event speciation. This definition of dispersal is not to be confused with definitions from ecology or population biology that concern the movement of individuals within a taxon's established range, seasonal migration, vagrant occurrences, etc.
- *Extinction*: An event involving the disappearance of an organism from part of its range (e.g., local extinction, extirpation, or range contraction) or throughout its geographical range (e.g. global extinction or lineage extinction). Extirpation in a widespread species

might only result in range contraction, whereas extirpation in an endemic species might result in lineage extinction.

- *Speciation*: A phylogenetic branching event that causes one species to become two species. Biogeographers are particularly interested in the spatial context of speciation. For example, speciation can be caused by a barrier interrupting the gene flow and isolating two populations, as in allopatric or geographic speciation; or by phenotypic, reproductive, or ecological divergence in two co-occurring populations that ultimately results in lineage divergence due to lack of gene flow, as in ecological speciation.
- *Vicariance and founder event speciation*: Vicariance is a historical scenario that involves the geographic range of an ancestral species dividing into two or more fragments due to climatic or geographic barrier formation (e.g., a new ocean basin opening), followed by allopatric speciation. Founder-event speciation is dispersal over a barrier followed by allopatric speciation. Because the final outcome is allopatric speciation in both cases, the timing of events is often used to infer which scenario took place. If the barrier and speciation event are similar in age, then vicariance is the preferred explanation. If the barrier is older than the speciation event, then founder-event speciation is preferred.

Biogeography Articles in Systematic Biology and Systematic Zoology

We analyzed the dataset of nearly all *Systematic Biology* and *Systematic Zoology* articles published between 1952 and July 2025 curated by (Landis & Donoghue, 2026). For each year, we collected all scientific publications (research articles, points of view, book reviews, spotlight articles, software articles, and symposium articles), and then counted both the total number of articles and the number of articles whose titles and/or abstracts contained any of the following substrings (partial matches allowed): *‘biogeo’, ‘geograph’, ‘vicarian’, ‘allopatr’, ‘sympatr’, ‘coloniz’, ‘disperse’, ‘dispersal’.* This procedure recovered 3,768 scientific articles, with 621 articles containing biogeography terms (16.3%). We then computed the proportion of biogeography-associated articles published per year (Appendix Figure 1). The shift from lower to higher proportions around 1970 is potentially caused by the fact that almost no *Systematic Biology/Zoology* articles published before 1967 had abstracts, so only (short) titles for those articles offered opportunities to associate them with biogeography terms. However, the bump may also coincide with the rising interest in biogeography among systematists (and cladists, in particular) at that time. Volume 37 from 1988 contained an unusually high proportion of biogeography-associated articles (52.9%). Two of the four issues published that year (Issues 3 and 4) are associated with the Society of Systematic Biologists’ "Symposium on vicariance biogeography: Theory, methods, and applications", whose proceedings are dedicated to the memory of the eminent biogeographer, Donn Rosen (1929-1986). All articles in those issues concern biogeography.



Appendix Figure 1. The number (left) and proportion (right) of scientific articles whose titles and/or abstracts contain biogeography terms that were published in *Systematic Biology* or *Systematic Zoology* during 1967 to 2025. See Appendix for details.

References

Albert, J.S., Schoolmaster Jr, D.R., Tagliacollo, V., & Duke-Sylvester, S.M., (2017). Barrier displacement on a neutral landscape: toward a theory of continental biogeography. *Systematic Biology*, 66(2), 167–182.

Alroy, J., Marshall, C., Bambach, R., Bezusko, K., Foote, M., Fürsich, F., Hansen, T., Holland, S., Ivany, L., Jablonski, D., & others. (2001). Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proceedings of the National Academy of Sciences*, *98*(11), 6261–6266.

Alves, J. M., Prado-Lopez, S., Cameselle-Teijeiro, J. M., & Posada, D. (2019). Rapid evolution and biogeographic spread in a colorectal cancer. *Nature Communications*, *10*(1), 5139.

Arias, J. S. (2024). Phylogenetic biogeography inference using dynamic paleogeography models and explicit geographic ranges. *Systematic Biology*, *73*(6), 995–1014.

Avise, J. C. (2000). *Phylogeography: The history and formation of species*. Harvard university press.

Barreda, V.D., Palazzesi, L., Telleria, M.C., Olivero, E.B., Raine, J.I. and Forest, F., 2015. Early evolution of the angiosperm clade Asteraceae in the Cretaceous of Antarctica. Proceedings of the National Academy of Sciences, 112(35), pp.10989-10994.

Bedford, T., Greninger, A. L., Roychoudhury, P., Starita, L. M., Famulare, M., Huang, M.-L., Nalla, A., Pepper, G., Reinhardt, A., Xie, H., & others. (2020). Cryptic transmission of SARS-CoV-2 in Washington state. *Science*, *370*(6516), 571–575.

Bedford, T., Riley, S., Barr, I. G., Broor, S., Chadha, M., Cox, N. J., Daniels, R. S., Gunasekaran, C. P., Hurt, A. C., Kelso, A., & others. (2015). Global circulation patterns of seasonal influenza viruses vary with antigenic drift. *Nature*, *523*(7559), 217–220.

Benson, R.B., Butler, R., Close, R.A., Saupe, E. and Rabosky, D.L., 2021. Biodiversity across space and time in the fossil record. Current Biology, 31(19), pp.R1225-R1236.

Bielejec, F., Lemey, P., Baele, G., Rambaut, A., & Suchard, M. A. (2014). Inferring heterogeneous evolutionary processes through time: From sequence substitution to phylogeography. *Systematic Biology*, *63*(4), 493–504.

Bilofsky, H. S., Burks, C., Fickett, J. W., Goad, W. B., Lewitter, F. I., Rindone, W. P., Swindell, C. D., & Tung, C.-S. (1986). The GenBank genetic sequence databank. *Nucleic Acids Research*, *14*(1), 1–4.

Bokma, F. (2002). Detection of punctuated equilibrium from molecular phylogenies. *Journal of Evolutionary Biology*, *15*(6), 1048–1056.

Brooks, D. R., & McLennan, D. A. (1991). *Phylogeny, ecology, and behavior: A research program in comparative biology*. University of Chicago press.

Buerki, S., Forest, F., Alvarez, N., Nylander, J.A.A., Arrigo, N. & Sanmartín, I. (2011). An evaluation of new parsimony-based versus parametric inference methods in biogeography: a case study using the globally distributed plant family Sapindaceae. *Journal of Biogeography*, *38*, *531-550*.

Brundin, L. (1966). Transantarctic relationships and their significance, as evidenced by chironomid midges. *Kungliga Svenska Vetenskapsakademiens Handlingar*, *11*(1), 1–472.

Caetano, D. S., O'Meara, B. C., & Beaulieu, J. M. (2018). Hidden state models improve state-dependent diversification approaches, including biogeographical models: HMM AND THE ADEQUACY OF SSE MODELS. *Evolution*, *72*(11), 2308–2324.

Cardillo, M., Weston, P. H., Reynolds, Z. K. M., Olde, P. M., Mast, A. R., Lemmon, E. M., Lemmon, A. R., & Bromham, L. (2017). The phylogeny and biogeography of *Hakea* (Proteaceae) reveals the role of biome shifts in a continental plant radiation. *Evolution*, *71*(8), 1928–1943.

Charleston, M. (1998). Jungles: A new solution to the host/parasite phylogeny reconciliation problem. *Mathematical Biosciences*, *149*(2), 191–223.

Chroni, A., Miura, S., Oladeinde, O., Aly, V., & Kumar, S. (2021). Migrations of cancer cells through the lens of phylogenetic biogeography. *Scientific Reports*, *11*(1), 17184.

Cook, L. G., & Crisp, M. D. (2005). Not so ancient: The extant crown group of *Nothofagus* represents a post-Gondwanan radiation. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1580), 2535–2544.

Coop, G., Pickrell, J.K., Novembre, J., Kudaravalli, S., Li, J., Absher, D., Myers, R.M., Cavalli-Sforza, L.L., Feldman, M.W. and Pritchard, J.K., 2009. The role of geography in human adaptation. PLoS genetics, 5(6), p.e1000500.

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

Coyne, J., & Orr, H. A. (2004). *Speciation*. Sinaeur.

Cox CB, Moore PD. Biogeography: an ecological and evolutionary approach. 8th ed. Hoboken, NJ: Wiley; 2010

Cracraft, J. (1972). Continental drift and Australian bird biogeography. *The Emu*, *72*, 171-174.

Cracraft, J. (1973). Continental drift, paleoclimatology, and the evolution and biogeography of birds. *J. Zoological Society London*, *169*, *455-545*.

Cracraft, J. (1974). Phylogeny and evolution of the ratite birds. *Ibis*, *116*(4), 494-521.

Crisp, M. D., & Cook, L. G. (2005). Do early branching lineages signify ancestral traits? *Trends in Ecology & Evolution*, *20*(3), 122–128.

Crisp, M. D., Trewick, S. A., & Cook, L. G. (2011). Hypothesis testing in biogeography. *Trends in Ecology & Evolution*, *26*(2), 66–72.

Croizat, L. (1958). *Panbiogeography*. [published by the author].

Croizat, L., Nelson, G., & Rosen, D. E. (1974). Centers of Origin and Related Concepts. *Systematic Zoology*, *23*(2), 265–287.

Cybis, G. B., Sinsheimer, J. S., Lemey, P., & Suchard, M. A. (2013). Graph hierarchies for phylogeography. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1614), 20120206.

Czekanski-Moir, J. E., & Rundell, R. J. (2019). The Ecology of Nonecological Speciation and Nonadaptive Radiations. *Trends in Ecology & Evolution*, *34*(5), 400–415.

Darlington, P. J. (1959). Area, Climate, and Evolution. *Evolution*, *13*(4), 488–510.

Darwin, C., & Wallace, A. R. (1858). *On the tendency of species to form varieties: And on the perpetuation of varieties and species by natural means of selection*.

de la Peña, P.G., Iglesias, G., Talavera, E., Meseguer, A.S., & Sanmartín, I. (in press). On the utility of Deep Learning for model classification and parameter estimation on complex diversification scenarios. *Systematic Biology*, doi: 10.1093/sysbio/syag030.

de Queiroz, A. (2014). *The monkey's voyage: How improbable journeys shaped the history of life*. Basic Books (AZ).

Ding, W.N., Ree, R.H., Spicer, R.A. and Xing, Y.W., 2020. Ancient orogenic and monsoon-driven assembly of the world's richest temperate alpine flora. Science, 369(6503), pp.578-581.

Dobzhansky, T. (1937). *Genetics and the Origin of Species*. Columbia University Press.

Donoghue, M. J., & Moore, B. R. (2003). Toward an Integrative Historical Biogeography1. *Integrative and Comparative Biology*, *43*(2), 261–270.

Donoghue, M.J. and Edwards, E.J., 2014. Biome shifts and niche evolution in plants. Annual Review of Ecology, Evolution, and Systematics, 45(1), pp.547-572.

Du Toit, A. L. (1937). *Our wandering continents: An hypothesis of continental drifting*. Oliver & Boyd.

Du Toit, A. L. (1944). Tertiary mammals and continental drift; a rejoinder to George G. Simpson. *American Journal of Science*, *242*(3), 145–163.

Du Toit, A. L., & Reed, F. R. C. (1927). *A geological comparison of South America with South Africa*. Carnegie Institution of Washington.

Dupin, J., Matzke, N. J., Särkinen, T., Knapp, S., Olmstead, R. G., Bohs, L., & Smith, S. D. (2017). Bayesian estimation of the global biogeographical history of the Solanaceae. *Journal of Biogeography*, *44*(4), 887–899.

Ebach, M. C. (2015). Origins of biogeography. *History, Philosophy and Theory of the Life Sciences*, *13*, 1–20.

Ebach, M. C., & Humphries, C. J. (2002). Cladistic biogeography and the art of discovery. In *Journal of Biogeography* (Vol. 29, Issue 4, pp. 427–444). Wiley Online Library.

Ebach, M. C., Humphries, C. J., & Williams, D. M. (2003). Phylogenetic biogeography deconstructed. *Journal of Biogeography*, *30*(9), 1285–1296.

Edwards, A. W. F & Cavalli-Sforza, L. L. (1964). Reconstruction of evolutionary trees. pp. 67-76 in

Phenetic and Phylogenetic Classification, ed. V. H. Heywood and J. McNeill. Systematics Association Publ. No. 6, London.

Faria, N. R., Suchard, M. A., Rambaut, A., Streicker, D. G., & Lemey, P. (2013). Simultaneously reconstructing viral cross-species transmission history and identifying the underlying constraints. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1614), 20120196.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, *17*(6), 368–376.

Garcia-R, J.C. and Matzke, N.J., 2021. Trait-dependent dispersal in rails (Aves: Rallidae): Historical biogeography of a cosmopolitan bird clade. Molecular Phylogenetics and Evolution, 159, p.107106.

Global Biodiversity Information Facility (2001) GBIF Memorandum of Understanding. Available at: https://www.gbif.org/document/80661/gbif-memorandum-of-understanding

Goldberg, E. E., Lancaster, L. T., & Ree, R. H. (2011). Phylogenetic Inference of Reciprocal Effects between Geographic Range Evolution and Diversification. *Systematic Biology*, *60*(4), 451–465.

Griffin, C. T., Wynd, B. M., Munyikwa, D., Broderick, T. J., Zondo, M., Tolan, S., Langer, M. C., Nesbitt, S. J., & Taruvinga, H. R. (2022). Africa’s oldest dinosaurs reveal early suppression of dinosaur distribution. *Nature*, *609*(7926), 313–319.

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

Heads, M. (2005). Dating nodes on molecular phylogenies: A critique of molecular biogeography. *Cladistics*, *21*(1), 62–78.

Heads, M. (2012). Bayesian transmogrification of clade divergence dates: a critique. *Journal of Biogeography*, *39*(10), 1749-1756.

Hennig, W. (1966). *Phylogenetic systematics*. University of Illinois Press.

Hess, H. H. (1962). History of ocean basins. In *Petrologic Studies: A volume in honor of A. F. Buddington* (pp. 599–620). Geological Society of America.

Hoberg, E. P., & Brooks, D. R. (2008). A macroevolutionary mosaic: Episodic host-switching, geographical colonization and diversification in complex host–parasite systems. *Journal of Biogeography*, *35*(9), 1533–1550.

Holmes, A. (1944). *Principles of Physical Geology*. Thomas Nelson & Sons.

Hooker, J. D. (1853). *Introductory Essay to the Flora of New Zealand*. Lovell Reeve.

Hortal, J., De Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. and Ladle, R.J. (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics, 46(1), 523-549*.

Huelsenbeck, J. P., & Imennov, N. S. (2002). Geographic origin of human mitochondrial DNA: accommodating phylogenetic uncertainty and model comparison. *Systematic Biology*, *51*(1), 155–165.

Hull, D. L. (1990). *Science as a process: An evolutionary account of the social and conceptual development of science*. University of Chicago Press.

Humphries, C. J., & Parenti, L. R. (1999). *Cladistic biogeography*. OUP Oxford.

Huxley, J. (1942). *Evolution: The Modern Synthesis*. George Allen & Unwin.

Knowles, L. L., & Maddison, W. P. (2002). Statistical phylogeography. *Molecular Ecology*, *11*(12).

Klaus, K. & Matzke, N. J. (2020). Statistical Comparison of Trait-Dependent Biogeographical Models Indicates That Podocarpaceae Dispersal Is Influenced by Both Seed Cone Traits and Geographical Distance. *Systematic Biology*, *69(1)*, *61–75*.

Kodandaramaiah, U. (2011). Tectonic calibrations in molecular dating. *Current Zoology*, *57*(1), 116–124.

Kossmat, F. (1936). *Paläogeographie und Tektonik*. B.G. Teubner.

Landis, M. J. (2017). Biogeographic dating of speciation times using paleogeographically informed processes. Systematic Biology, 66(2), 128-144.

Landis, M. J., & Donoghue, M. J. (2026). Seventy-Five Years of Systematic Biology: Looking Back, Moving Forward. *Systematic Biology*, *75*(1), 3-13.

Landis, M. J., Edwards, E. J., & Donoghue, M. J. (2021a). Modeling Phylogenetic Biome Shifts on a Planet with a Past. *Systematic Biology*, *70*(1), 86–107.

Landis, M. J., Eaton, D. A. R., Clement, W. L., Park, B., Spriggs, E. L., Sweeney, P. W., Edwards, E. J., & Donoghue, M. J. (2021b). Joint Phylogenetic Estimation of Geographic Movements and Biome Shifts during the Global Diversification of *Viburnum*. *Systematic Biology*, *70*(1), 67–85.

Landis, M. J., Freyman, W. A., & Baldwin, B. G. (2018). Retracing the Hawaiian silversword radiation despite phylogenetic, biogeographic, and paleogeographic uncertainty. *Evolution*, *72*(11), 2343–2359.

Landis, M. J., Quintero, I., Muñoz, M. M., Zapata, F., & Donoghue, M. J. (2022). Phylogenetic inference of where species spread or split across barriers. *Proceedings of the National*

Academy of Sciences, *119*(13), e2116948119.

Landis, M. J., & Thompson, A. (2025). phyddle: software for exploring phylogenetic models with deep learning. Systematic Biology, 74(6), 1007-1019.

Lambert, S., Voznica, J., & Morlon, H. (2023). Deep learning from phylogenies for diversification analyses. Systematic Biology, 72(6), 1262-1279.

Lemey, P., Rambaut, A., Drummond, A. J., & Suchard, M. A. (2009). Bayesian phylogeography finds its roots. *PLoS Computational Biology*, *5*(9), e1000520.

Lemey, P., Rambaut, A., Welch, J. J., & Suchard, M. A. (2010). Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology and Evolution*, *27*(8), 1877–1885.

Lemmon, A. R., & Lemmon, E. M. (2008). A likelihood framework for estimating phylogeographic history on a continuous landscape. *Systematic Biology*, *57*(4), 544–561.

Lewinsohn, M.A., Bedford, T., Müller, N.F., & Feder, A. F. (2023). State-dependent evolutionary models reveal modes of solid tumour growth. *Nature Ecology & Evolution*, *7*, *581–596*.

Lieberman, B. S. (2002). Phylogenetic biogeography with and without the fossil record: Gauging the effects of extinction and paleontological incompleteness. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *178*(1–2), 39–52.

Louca, S. (2021). Phylogeographic estimation and simulation of global diffusive dispersal. *Systematic Biology*, *70*(2), 340–359.

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

Maddison, W. P., Midford, P. E., & Otto, S. P. (2007). Estimating a Binary Character’s Effect on

Speciation and Extinction. *Systematic Biology*, *56*(5), 701–710.

Mao, K., Milne, R. I., Zhang, L., Peng, Y., Liu, J., Thomas, P., Mill, R. R., & S. Renner, S. (2012). Distribution of living Cupressaceae reflects the breakup of Pangea. *Proceedings of the National Academy of Sciences*, *109*(20), 7793–7798.

Marshall, C., & Liebherr, J. (2000). Cladistic biogeography of the Mexican transition zone. *Journal of Biogeography*, *27*(1), 203–216.

Massana, K. A., Beaulieu, J. M., Matzke, N. J., & O'Meara, B. C. (2015). Non-null effects of the null range in biogeographic models: Exploring parameter estimation in the DEC model. *bioRxiv*, 026914.

Matzke, N. J. (2014). Model Selection in Historical Biogeography Reveals that Founder-Event Speciation Is a Crucial Process in Island Clades. *Systematic Biology*, *63*(6), 951–970.

Matzke, N. J. (2022). Statistical comparison of DEC and DEC+ J is identical to comparison of two ClaSSE submodels, and is therefore valid. *Journal of Biogeography*, *49*(10), 1805–1824.

Mayr, E. (1942). *Systematics and the origin of species*. Harvard University Press.

Mayr, E. (1954). Change of genetic environment and evolution. In *Evolution as process* (pp. 157–180). Macmilliam.

Mayr, E., & Provine, W. B. (1980). *The evolutionary synthesis: Perspectives on the unification of biology*. Harvard University Press.

Mayr, G., 2004. Old World fossil record of modern-type hummingbirds. Science, 304(5672), pp.861-864.

McKenna, M. C. (1973). Sweepstakes, filters, corridors, Noah's arks, and beached Viking funeral ships in palaeogeography. In *Implications of Continental Drift to the Earth Sciences* (pp.

293–308). Academic Press.

Meseguer, A. S., Lobo, J. M., Ree, R., Beerling, D. J., & Sanmartín, I. (2015). Integrating Fossils, Phylogenies, and Niche Models into Biogeography to Reveal Ancient Evolutionary History: The Case of Hypericum (Hypericaceae). *Systematic Biology*, *64*(2), 215–232.

Merckx, V.S., Hendriks, K.P., Beentjes, K.K., Mennes, C.B., Becking, L.E., Peijnenburg, K.T., Afendy, A., Arumugam, N., de Boer, H., Biun, A. and Buang, M.M., 2015. Evolution of endemism on a young tropical mountain. Nature, 524(7565), pp.347-350.

Morrone, J. J. (2021). Croizat's dangerous ideas: Practices, prejudices, and politics in contemporary biogeography. *History and Philosophy of the Life Sciences*, *43*(2), 77.

Morrone, J.J., 2022. Matthew's (1915) climate and evolution, the "New York School of Biogeography", and the rise and fall of "Holarcticism". History and Philosophy of the Life Sciences, 44(2), p.15.

Nadeau, S. A., Vaughan, T. G., Scire, J., Huisman, J. S., & Stadler, T. (2021). The origin and early spread of SARS-CoV-2 in Europe. *Proceedings of the National Academy of Sciences*, *118*(9), e2012008118.

Nagel, A. S. & Landis M. J. (2026). Ancestral state reconstruction with discrete characters using deep learning. bioRxiv 2026.03.19.712918.

Nelson, G. (1974). Historical Biogeography: An Alternative Formalization. *Systematic Zoology*, *23(4)*, 555–558.

Nelson, G. (1978). From Candolle to Croizat: Comments on the history of biogeography. *Journal of the History of Biology*, *11*, 269–305.

Nelson, G., & Ladiges, P. Y. (2001). Gondwana, vicariance biogeography and the New York

School revisited. *Australian Journal of Botany*, *49*(3), 389–409.

Nelson, G., & Platnick, N. I. (1981). *Systematics and Biogeography: Cladistics and Vicariance*. Columbia University Press.

Nepokroeff, M., Sytsma, K. J., Wagner, W. L., & Zimmer, E. A. (2003). Reconstructing ancestral patterns of colonization and dispersal in the Hawaiian understory tree genus Psychotria (Rubiaceae): A comparison of parsimony and likelihood approaches. *Systematic Biology*, *52*(6), 820–838.

Nielsen, R. (2002). Mapping mutations on phylogenies. *Systematic Biology*, *51*(5), 729–739.

Nylander, J. A., Olsson, U., Alström, P., & Sanmartín, I. (2008). Accounting for phylogenetic uncertainty in biogeography: A Bayesian approach to dispersal-vicariance analysis of the thrushes (Aves: Turdus). *Systematic Biology*, *57*(2), 257–268.

Nylinder, S., Lemey, P., De Bruyn, M., Suchard, M. A., Pfeil, B. E., Walsh, N., & Anderberg, A. A. (2014). On the biogeography of Centipeda: A species-tree diffusion approach. *Systematic Biology*, *63*(2), 178–191.

Overcast, I., Ruffley, M., Rosindell, J., Harmon, L., Borges, P.A., Emerson, B.C., Etienne, R.S., Gillespie, R., Krehenwinkel, H., Mahler, D.L., Massol, F., Parent, C. E., Patiño, J., Peter, B., Week, B., Wagner, C., Hickerson, M. J., Rominger, A. (2021). A unified model of species abundance, genetic diversity, and functional diversity reveals the mechanisms structuring ecological communities. *Molecular Ecology Resources*, 21(8), 2782–2800.

Page, R. D. (2003). *Tangled trees: Phylogeny, cospeciation, and coevolution*. University of Chicago press.

Page, R. D. M. (1990). Temporal Congruence and Cladistic Analysis of Biogeography and

Cospeciation. *Systematic Biology*, *39*(3), 205–226.

Page, R. D. M. (1994). Maps Between Trees and Cladistic Analysis of Historical Associations among Genes,Organisms, and Areas. *Systematic Biology*, *43*(1), 58–77.

Pagel, M. (1999). The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic Biology*, *48*(3), 612–622.

Parenti, L. R., & Humphries, C. J. (2004). Historical biogeography, the natural science. *Taxon*.

Platnick, N. I. (1977). Cladograms, phylogenetic trees, and hypothesis testing. *Systematic Zoology*, *26*(4), 438–442.

Platnick, N. I., & Nelson, G. (1978). A Method of Analysis for Historical Biogeography. *Systematic Zoology*, *27*(1), 1.

Qin, T., van Benthem, K. J., Valente, L., & Etienne, R. S. (2026). Parameter estimation from phylogenetic trees using neural networks and ensemble learning. *Systematic Biology*, 75(2), 344-365.

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

Quintero, I., Landis, M. J., Jetz, W., & Morlon, H. (2023). The build-up of the present-day tropical diversity of tetrapods. *Proceedings of the National Academy of Sciences*, *120*(20), e2220672120.

Raven, P. H. & D. I. Axelrod. (1972). Plate tectonics and Australasian paleobiogeography. *Science, 176, 1379–1386.*

Raven, P. H., & Axelrod, D. I. (1974). Angiosperm biogeography and past continental movements. *Annals of the Missouri Botanical Garden*, *61*(3), 539–673.

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

Ree, R. H., & Sanmartín, I. (2009). Prospects and challenges for parametric models in historical biogeographical inference. *Journal of Biogeography*, *36*(7), 1211–1220.

Ree, R. H., & Sanmartín, I. (2018). Conceptual and statistical problems with the DEC +J model of founder-event speciation and its comparison with DEC via model selection. *Journal of Biogeography*, *45*(4), 741–749.

Ree, R. H., & Smith, S. A. (2008). Maximum Likelihood Inference of Geographic Range Evolution by Dispersal, Local Extinction, and Cladogenesis. *Systematic Biology*, *57*(1), 4–14.

Ronquist, F. (1997). Dispersal-Vicariance Analysis: A New Approach to the Quantification of Historical Biogeography. *Systematic Biology*, *46*(1), 195–203.

Ronquist, F. (2003). Parsimony analysis of coevolving species associations. *Tangled Trees: Phylogeny, Cospeciation and Coevolution*, 22–64.

Ronquist, F. (2004). Bayesian inference of character evolution. *Trends in Ecology & Evolution*, *19*(9), 475–481.

Ronquist, F., & Sanmartín, I. (2011). Phylogenetic Methods in Biogeography. *Annual Review of Ecology, Evolution, and Systematics*, *42*(1), 441–464.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, *74*(12), 5463–5467.

Sanmartín, I. & F. Ronquist, (2004). Southern Hemisphere Biogeography Inferred by Event-Based Models: Plant versus Animal Patterns. *Systematic Biology*, 53(2), 278–298.

Sanmartín, I. (2007). Event-based biogeography: Integrating patterns, processes, and time. In *Biogeography in a Changing World*. CRC Press.

Sanmartín, I. (2012). Historical biogeography: Evolution in time and space. *Evolution: Education and Outreach*, *5*(4), 555–568.

Sanmartin, I. (2022). Analytical approaches in biogeography: advances and challenges. In *Biogeography: An Integrative Approach of the Evolution of Living*. Eric Guilbert. ed. John Wiley & Sons, Inc. 25-52.

Sanmartín, I., Enghoff, H., & Ronquist, F. (2001). Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biological Journal of the Linnean Society*, *73*(4), 345–390.

Sanmartín, I., Anderson, C. L., Alarcon, M., Ronquist, F., & Aldasoro, J. J. (2010). Bayesian island biogeography in a continental setting: the Rand Flora case. *Biology letters*, 6(5), 703-707.

Skeels, A., & Cardillo, M. (2019). Reconstructing the geography of speciation from contemporary biodiversity data. *The American Naturalist*, 193(2), 240-255.

Silvestro, D., Zizka, A., Bacon, C. D., Cascales-Minana, B., Salamin, N., & Antonelli, A. (2016). Fossil biogeography: a new model to infer dispersal, extinction and sampling from palaeontological data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1691).

Simpson, G. G. (1943). Mammals and the nature of continents. *American Journal of Science*, *241*(1), 1–31.

Soewongsono, A. C., & Landis, M. J. (2026). Phylogenetic estimation of diversity-dependent biogeographic rates using deep learning. bioRxiv, 2026-02.

Sukumaran, J., Economo, E. P., & Lacey Knowles, L. (2016). Machine Learning Biogeographic Processes from Biotic Patterns: A New Trait-Dependent Dispersal and Diversification Model with Model Choice By Simulation-Trained Discriminant Analysis. *Systematic Biology*, *65*(3), 525–545.

Sukumaran, J. and Knowles, L.L., 2018. Trait-dependent biogeography:(re) integrating biology into probabilistic historical biogeographical models. Trends in Ecology & Evolution, 33(6), pp.390-398.

Swiston, S. K., & Landis, M. J. (2025). Testing relationships between multiple regional features and biogeographic processes of speciation, extinction, and dispersal. *Systematic Biology*, *74*(2), 282–300.

Thompson, A., Liebeskind, B. J., Scully, E. J., & Landis, M. J. Deep Learning and Likelihood Approaches for Viral Phylogeography Converge on the Same Answers Whether the Inference Model Is Right or Wrong, *Systematic Biology*, Volume 73, Issue 1, January 2024, Pages 183–206.

Thorne, J. L., & Kishino, H. (2002). Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology*, *51*(5), 689–702.

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

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

Van Veller, M. G., Brooks, D., & Zandee, M. (2003). Cladistic and phylogenetic biogeography: The art and the science of discovery. In *Journal of Biogeography* (Vol. 30, Issue 3, pp. 319–329). Wiley Online Library.

Vasconcelos, T., Boyko, J.D. and Beaulieu, J.M., 2023. Linking mode of seed dispersal and climatic niche evolution in flowering plants. Journal of Biogeography, 50(1), pp.43-56.

Vine, F. J., & Matthews, D. H. (1963). Magnetic Anomalies Over Oceanic Ridges. *Nature*, *199*(4847), 947–949.

Voznica, J., Zhukova, A., Boskova, V., Saulnier, E., Lemoine, F., Moslonka-Lefebvre, M., & Gascuel, O. (2022). Deep learning from phylogenies to uncover the epidemiological dynamics of outbreaks. Nature Communications, 13(1), 3896.

Wegener, A. (1912). Die entstehung der kontinente. *Geologische Rundschau*, *3*(4), 276–292.

Wisniewski, A.L., Lloyd, G.T. and Slater, G.J., 2022. Extant species fail to estimate ancestral geographical ranges at older nodes in primate phylogeny. Proceedings of the Royal Society B, 289(1975), p.20212535.

Wojcicki, M., & Brooks, D. R. (2005). PACT: an efficient and powerful algorithm for generating area cladograms. *Journal of Biogeography*, *32*(5), 755–774.

Wood, H. M., Matzke, N. J., Gillespie, R. G., & Griswold, C. E. (2013). Treating fossils as terminal taxa in divergence time estimation reveals ancient vicariance patterns in the palpimanoid spiders. *Systematic Biology*, *62*(2), 264–284.

Wootton, L.M., Boucher, F.C., Pouchon, C., Roquet, C., Coissac, E., Renaud, J., Alsos, I.G., Valla, P.G., Husson, L., Bernet, M. and Perrier, C., 2025. The late rise of sky-island vegetation in the European Alps. Nature Plants, pp.1-12.

Worobey, M., Watts, T. D., McKay, R. A., Suchard, M. A., Granade, T., Teuwen, D. E., Koblin, B. A., Heneine, W., Lemey, P., & Jaffe, H. W. (2016). 1970s and 'Patient 0'HIV-1 genomes illuminate early HIV/AIDS history in North America. *Nature*, *539*(7627), 98–101.

Ye, J.W., Yang, Z.Z. and Tian, B., 2023. Tempo-spatial evolution of seed plant endemism in Taiwan island. Journal of Biogeography, 50(11), pp.1981-1991.

Zuckerkandl, E., & Pauling, L. (1962). Molecular disease, evolution, and genic heterogeneity. In *Horizons in Biochemistry* (pp. 189–225). Academic Press.