Botany and Geogenomics: constraining geological hypotheses with large-scale genetic data derived from plants

Ana M. Bedoya

Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, 70803, USA. Email: abedoya@lsu.edu; https://orcid.org/0000-0002-6245-3481

Author for correspondence:

Ana María Bedoya

Tel: +1(206)310-4972

Email: abedoya@lsu.edu

Abstract

Past studies in plant phylogenetics have shed light on how the geological history of our planet shaped plant evolution by establishing well-known patterns (e.g., how mountain uplift resulted in high rates of diversification and replicate radiations in montane plant taxa). Under this approach, information is transferred from geology to botany, by interpreting data in light of geological processes. In this synthesis, I propose a conceptual shift in this traditional approach to specifically transfer information from botany to geology. This conceptual shift follows the goals of the emerging field of geogenomics and emphasizes that plant phylogenetics can go beyond investigating patterns in light of landscape change, to reduce the inherent uncertainty in models of paleotopography, river system structure, and land connections through time. Current challenges that are specific to analytical approaches for plant geogenomics are discussed. I describe the scale at which various geological questions can be addressed from biological data, and what makes some groups of plants excellent model systems for geogenomics research. This synthesis highlights the critical role of classical botanical knowledge in identifying good study systems to unveil long-standing questions on how the earth evolved with the use of plant DNA.

Key words: admixture, biogeography, demographic modeling, geogenomics, geology, orogeny, phylogenetic networks, plant phylogenomics

Geological history has played a key role in shaping plant evolution (Gentry, 1982; Hughes et al., 2013; Sundaram et al., 2019). Geomorphic processes have resulted, among others, in the formation of islands and elevational gradients (Humboldt and Bonpland, 1807; Clague and Sherrod, 2014; Rahbek et al., 2019), leading to identifiable evolutionary patterns. These patterns include accelerated diversification (e.g., explosive radiations in Andean bellflowers and *Lupinus;* Hughes and Eastwood, 2006; Lagomarsino et al., 2016), convergent evolution (e.g., evolution of reproductive traits in Hawaiian Lobeliads and of life forms in temperate and arctic-alpine *Androsace*; Givnish et al., 2009; Roquet et al., 2013), limited gene flow across landscape units (e.g., in *Amphirrox longifolia* and two species of *Marathrum* across the Rio Negro and Andean mountains respectively; Nazareno et al., 2017; Bedoya et al., 2021), and replicated radiations (e.g., repeated evolution of leaf ecomorphs in a Neotropical clade of *Viburnum*; Donoghue et al., 2022). Plant phylogenetics research has shed light on such patterns by interpreting biological data in reference to geological processes. Here, I propose a shift in this traditional approach, emphasizing that plant phylogenetics can go beyond investigating patterns in light of landscape change, to specifically transfer information from botany to geology (Bedoya et al., 2021).

The use of patterns inferred from large-scale genetic data to constrain and reduce uncertainty in geological hypotheses, was defined as the goal of geogenomics by Baker et al. (2014). This emerging field involves the reciprocal integration of geologic, climatic, and genomic data to test landscape evolution hypotheses at various time scales from biological data (Dolby et al., 2022). Here I describe how by integrating natural history, phylogenetics, and population genetics, botanical research can inform geological processes. In this synthesis, I also contribute practical considerations to geogenomics that are specific to plant research. I review studies that, by introducing promising systems to constrain geological hypotheses and informing current models of defining geological and climatic events, exemplify my description on how to shift the paradigm in plant phylogenetics.

WHY AND HOW TO SHIFT THE PARADIGM?

Plant evolutionary patterns can inform paleogeograhic dynamics— Since the proposal of the 'riverine barrier' hypothesis by A.R. Wallace in 1852, historical biogeography has aimed to study distribution patterns of organisms with reference to landscape dynamics. Research in this area has tremendously advanced our understanding of plant evolution on our constantly changing planet. Examples where dated phylogenies were interpreted according to landscape model reconstructions abound (e.g., biogeographic studies in Asteraceae, Bromeliaceae, Lamiaceae, Malpighiaceae, Myristicaceae, Orchidaceae, and Solanaceae (Davis et al., 2002; Bell, 2005; Givnish et al., 2011, 2016; Dupin et al., 2017; Frost et al., 2017, 2022; Mandel et al., 2019; Rose et al., 2022). A recent and more sophisticated approach to modeling species geographic ranges is biogeographic dating, a time-stratified framework where paleogeographic hypotheses impose constraints on the rates of biogeographic change and vice versa (Landis, 2017). This framework allows for the incorporation of uncertainty in estimates of formation and change of landscape units in the past. For example, proposed times of island formation in the Hawaiian archipelago (Lim and Marshall, 2017) were incorporated as the maximum ages at which dispersal to each island could have taken place in a biogeographic study of Hawaiian silverswords (Asteraceae; Landis et al., 2018; Fig. 1A). This analytical framework was expanded to re-examine the ancestral biome reconstruction of *Viburnum* (Adoxaceae; Landis et al., 2021).

A major assumption of these approaches is that there is a paleogeographic or climatic model of the context in which plants evolved in a given area. The example above is unusual in the extent to which we know exact dates for island formation in the Hawaiian archipelago. However, unifying, accurate, and precise landscape models, are unavailable for a majority of regions and across time scales. This lack of precise models can be explained by uncertainty in analytical tools for dating strata (Schoene et al., 2013), the complexity of past landscape dynamics, and the high degree of heterogeneity of topographic units. While geological advances continue to improve paleogeographic reconstructions, the growing amount of genomic data being generated in current times presents an opportunity to improve what we know about landscape configurations over time. A shift in the paradigm of using landscape models to inform biological processes to instead using biological data (i.e., large-scale genetic data) to constrain geological hypotheses (Fig 1B), is a powerful approach to explore long-standing questions in geology and plant evolution.

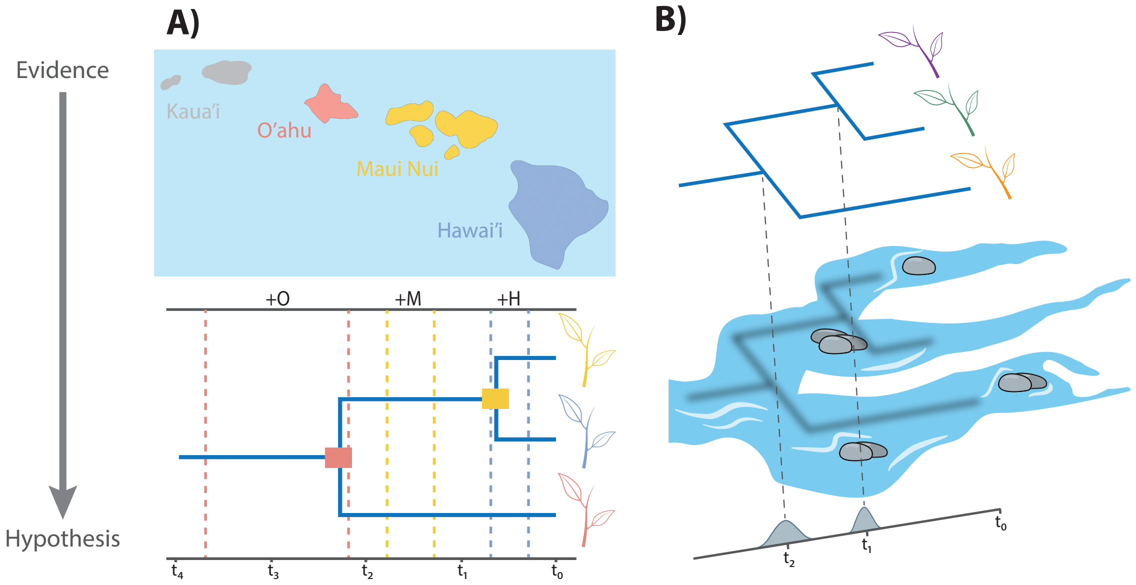

Figure 1. Approaches to questions at the intersection of geology and plant evolution from evidence to generation of a hypothesis. A) Biogeographic dating *sensu* Landis et al., 2017; 2018. Modeling of ancestral ranges is conducted by imposing proposed times for the formation of islands in the Hawaiian archipelago (top) as constraints on the rates of biogeographic change (bottom). B) Geogenomics approach, where divergence times for three populations of river plants are used to constrain the approximate timing and pattern of connections of paleodrainages through time.

Approaches to plant geogenomics— Genomic signatures of past geologic events may be identified with four main approaches: 1) estimation of divergence times of species/populations distributed across landscape units (Fig. 1B); 2) identification of the distribution of genetic variation across the landscape (i.e., genetic structure; Fig. 2A); 3) investigation of introgression events (e.g., phylogenetic network inference, summary methods for the detection of gene flow from site patterns, and likelihood-based methods for inference of introgression; Fig. 2B; Green et al., 2010; Durand et al., 2011; Blischak et al., 2018; Nauheimer et al., 2021; Ji et al., 2023); and 4) demographic modeling (i.e., using site frequency spectra and whole genomes; Fig. 2C; Adams and Hudson, 2004; Gutenkunst et al., 2009; Li and Durbin, 2011; Schiffels and Durbin, 2014).

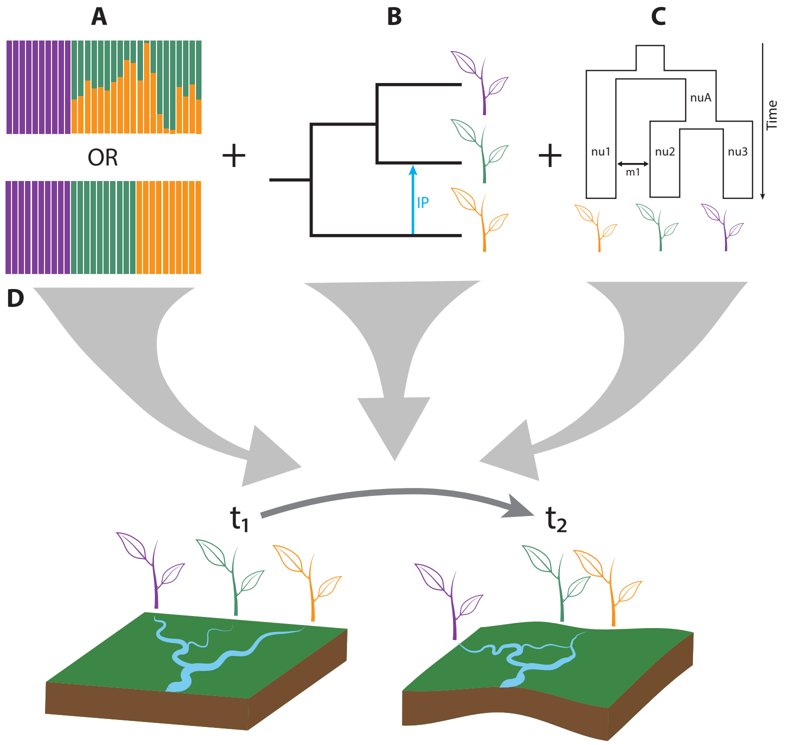

Figure 2. Toy example of three lines of evidence that can be used to identify a landscape change event from biological data. Patterns seen from A) inference of admixture proportions, B) identification of introgression events through phylogenetic network inference, and C) modeling of past demographic dynamics. Such patterns indicate carry a signal left by D) a river capture event that took place at time t_2. The event resulted in secondary contact and admixture of two previously isolated plant populations (orange and green). IP: Inheritance probability or the proportion of genes contributed by each parental population to a reticulation node, if gene trees are used for network inference.

The example in Fig. 2 shows a river capture event (i.e., the diversion of part of the course of a river into another river; Fig. 2D) taking place at time t_2, and the patterns that could be identified with the approaches mentioned above. The river capture event resulted in secondary contact and admixture of two previously isolated populations of river plants (green and orange), as identified in an admixture plot (Fig. 2A, top). However, signatures of admixture could be eroded by subsequent events such as genetic drift (Fig. 2A, bottom), particularly in small populations (e.g., plant populations in savanna pockets within the Amazon rain forest, sky islands like the páramos, or river-rapids). Exploration of non-bifurcating relationships through phylogenetic network inference could also detect the branches of the tree involved in the reticulation event (green and orange populations; Fig. 2B), as well as the inheritance probability (IP; the proportion of genes contributed by each parental population to a reticulation node, if gene trees are used for network inference; (Solís-Lemus et al., 2017). Ideally, the 3D demographic model in Fig. 2C would be identified as the best fit for genomic data generated from the three populations (purple, green and orange). Variations of this model would include changes in population size and asymmetric gene flow. Concordant patterns across approaches would facilitate interpretation of the resulting patterns as indicators of a specific past geomorphic process.

Constraining landscape and climatic hypotheses with the use of time-trees, admixture plots, phylogenetic networks, and inferred demographic models as suggested above requires the joint application of at least two of the four approaches for various reasons. For example, if patterns of admixture are identified, they could be a result of retention of ancestral polymorphisms from shared ancestry (Lawson et al., 2018). Also, the signatures of polymorphism in truly admixed populations are expected to be removed with time (Clark, 1997). There are also practical limits to existing algorithms: phylogenetic network inference is computationally feasible only for a small number of terminal taxa, does not identify the direction at which gene flow took place, and may suffer from non-identifiability as complexity of the model increases (Yu et al., 2014; Pardi and Scornavacca, 2015; Solís-Lemus and Ané, 2016; Solís-Lemus et al., 2017). In addition, demographic models inferred from genetic data have been developed for a limited number of populations (Farleigh et al., 2021), suffer from non-identifiability (Terhorst and Song, 2015), and require *a priori* specifications of parameters and demographic events (unknown in most empirical systems; Li and Durbin, 2011; Loog, 2021).

Special caveats specific but not exclusive to plant geogenomics include the scarcity of fossil evidence useful for calibration for many plant groups (Crane et al., 2004; Schenk, 2016). Other caveats derive from the prevalence of hybridization, polyploidization, and various mating systems in plants (Stebbins, 1950) and the difficulty in developing analytical tools that accommodate these complexities (Blischak et al., 2023). However, some standard genotyping tools that enable ploidy specification and algorithms for genotype likelihood estimation are currently available (Garrison and Marth, 2012; Serang et al., 2012; Blischak et al., 2018b; Gerard et al., 2018; Clark et al., 2019). Other challenges include the identification of paralogous and orthologous loci prior to phylogenetic inference (Yang and Smith, 2014; Johnson et al., 2016; Gardner et al., 2021; Morales-Briones et al., 2022; Freyman et al., 2023; Mendez-Reneau et al., 2023), scarcity of resources for analysis of shared ancestry in polyploids (but see Kolář, 2021; Shastry et al., 2021), and demographic modeling under various ploidy and mating systems scenarios (but see Roux and Pannell, 2015; Blischak et al., 2023; Roux et al., 2023). Simulation studies testing the effect of strategies and violations of model assumptions in plant geogenomics research (Schenk, 2016; Stift et al., 2019) are critical to better constrain and propose plausible geological scenarios from biological data in light of uncertainty.

Selection of study systems and the problem of scale— Figure 1B shows how inferred divergence times of river plants could be used to propose a hypothesis for fluvial connections in the past. The efficacy of this approach requires careful selection of model systems whose evolutionary histories reflect that of rivers, as organismal responses to geomorphic barriers are highly heterogeneous (Araya-Donoso et al., 2022; Rodriguez-Muñoz et al., 2022). A famous example of differing responses to a common event is the variable timing of dispersal in terrestrial organisms (i.e., plants, birds, mammals, reptiles, amphibians, and arthropods) across the isthmus of Panama (Bacon et al., 2015). The hypothesis-testing approach shown in Figure 1B would necessitate that the plants studied live strictly in riverine habitats, is naturally distributed across rivers, and has a low dispersal ability. Once taxa that meet these characteristics are found, interpretation of analyses derived from genomic data for such study systems should consider other aspects of their biology like population size, mating systems, life cycle, and ploidy.

Estimating the timing of geomorphic events from genomic data is scale-dependent and the limitations mentioned above determine the type of research questions that are feasible. The example shown in Figure 1B is an oversimplification of a deep time (millions of years), large-scale problem, where the connections of river drainages and not of single rivers are inferred. This is the scale at which research questions are approached by classical biogeography. At a smaller-scale, the genomic signature of changes in river connections (Fig. 2) could be erased by changes in population size, and by the possibly large effect of genetic drift in removing polymorphisms in river plants with small population sizes (Bedoya et al., 2021). Thus, river capture events may be reconstructed most reliably only if occurring at a shallow time scale. Similarly, estimating the timing of connections across high-elevation ecosystems based on genomic data might be limited to the most recent events. For example, páramo plants experienced demographic changes due to the contraction and expansion of high-elevation ecosystems during the Pleistocene glaciations (Flantua et al., 2019), the most recent of which are identifiable from genomic data (Nevado et al., 2018).

PLANT EVOLUTION IN GEOGENOMICS RESEARCH OR HOW ELSE ARE PLANTS SPECIAL

The proposal of the geogenomics neologism (Baker et al., 2014) included the identification of major geologic questions in the Neotropics to be addressed with biological data. These included the history of mountain building (e.g., Andean uplift), river reconfiguration (e.g, the Amazon river), and formation of bridges between terrestrial ecosystems (e.g, closure of the Isthmus of Panama). The ground work for many of these problems consists of a model of the geologic history of northern South America since the Paleogene by Hoorn et al., 2010, which provided a review of paleontological, sedimentological, ecological, and geological data for the region. Over the past decade, research using evidence derived from plants has helped constrain hypotheses laid out by Hoorn et al., 2010, by Baker et al., 2014, and by the several prior studies that they were based on.

Fine-scale uplift history of different segments and Cordilleras of the Northern Andes— Landscape models indicate that the uplift of the Andes (which in northern South America are split into the Western, Central, and Eastern Cordilleras) was an asynchronous process that started in the Late Cretaceous (Horton, 2018; Gianni et al., 2020). A current challenge is to generate a finer-scale geological model that reflects the asynchrony and spatial heterogeneity of the Andes, particularly for the Western and Central cordilleras (Sanín et al., 2022). A recent study by Sanín et al., 2022 tackled this challenge and revised the geologic history of uplift across the three mountain ranges that constitute the northern Andes, using genomic data generated for the *Geonoma undata–G. orbignyana* species complex (Arecaceae). The study integrated inference of genetic structure, phylogenetic relationships, divergence times, and demographic modeling, with information from past geological studies in the region (Velandia et al., 2005; Mora et al., 2008; Veloza et al., 2012; Anderson et al., 2016; Montes et al., 2019, 2021; Mora-Páez et al., 2019; Noriega-Londoño et al., 2020). They provide evidence for a scenario where mountain segments were already uplifted but disconnected until the Pliocene, when continuous cordilleras were formed. Phylogeographic breaks coincide with the location of strike-slip faults (i.e., vertical fractures where blocks have moved horizontally). The authors suggest that extensional and compressional features of the faults may represent initial negative topography that is subsequently inverted (Dewey et al., 1998; Veloza et al., 2012; Montes et al., 2019; Zapata et al., 2023). This resulted in the connection of previously isolated mountain segments during the Pliocene. Phylogeographic inference in the *Tococa-Azteca* plant-ant mutualism (Torres Jimenez et al., 2021) and biogeographic modeling in the Rubiaceae (Antonelli et al., 2009), further support the existence of lowland passes across the Andean mountains, and the dispersal through lowland portals in frogs, mammals, and non-avian reptiles (Rodriguez-Muñoz et al., 2022).

Drainage basin reconfiguration through time across the northern Andes— In addition to being a major driver of Neotropical diversification (Antonelli et al., 2009; Antonelli and Sanmartín, 2011;

Sklenář et al., 2011; Smith et al., 2014; Hoorn et al., 2018), Andean uplift reconfigured watersheds across South America (Albert et al., 2006; Hoorn et al., 2010; Ruokolainen et al., 2019). However, the configuration of paleodrainages through time remains largely unknown. To propose a hypotheses for past river connections in northern South America, I have used two species of strictly riverine plants (*Marathrum*, Podostemaceae) with broad distributions across the Andes as a model system. I used genome-scale data for populations from different drainage basins to infer population structure, phylogenetic networks, divergence dating, and population summary statistics (Bedoya et al., 2021). After establishing that gene flow is limited in populations across drainage basins, the timing and pattern of population divergence across river drainages was used as evidence to propose that drainage basins became separated at ~12 and ~4 Ma as a result of major pulses of Andean uplift (Gregory-Woodzicki, 2000; Garzione et al., 2008; Hoorn et al., 2010; Mora et al., 2010; Anderson et al., 2016; Boschman, 2021). This paleodrainage model was subsequently validated by a paleogeographic model of northern South America through the Miocene, built from geochronological, petrographic, and thermochronological data (Zapata et al., 2023).

Additional botanical studies have introduced promising study systems to constrain hypotheses for river reconfiguration in the past, including *Mauritia flexuosa* (Arecaceae), whose populations are inferred to have strong genetic structure across drainage basins in the Amazon (Sander et al., 2018). Another example is *Amphirrox longifolia* (Violaceae), a species with limited gene flow across broad portions of the Rio Negro in the Amazon basin (Nazareno et al., 2017). These plant examples complement studies in animals that are constrained to river drainages (Albert et al., 2006; Picq et al., 2014; Tagliacollo et al., 2015). Comparative phylogeographic studies at various geographic scales for these and other plant systems with putative limited dispersal ability could constitute powerful tools to generate hypotheses for river configurations in the past.

Neogene and Pleistocene marine incursions in the Neotropics—

Landscape change models developed from palynological, sedimentological, geochemical, and macrofossil data indicate that the Pebas Lake, a vast wetland-like system of > 1 million km^2, was in place in northern and central South America from ca. 23–10 Ma (Hoorn, 1993; Wesselingh et al., 2002; Hoorn et al., 2010; Linhares et al., 2017; Jaramillo, 2023). Isotopic and palynological evidence shows a marine influence on the Pebas lake, but conflicting interpretations of the Miocene depositional environment and extent of marine incursions at the time remain (Frailey et al., 1988; Hoorn, 1993, 2006; Räsänen et al., 1995; Webb, 1995; Vonhof et al., 2003; Latrubesse et al., 2007, 2010; Jaramillo et al., 2017).

Based on distributional information, dispersal biology, and calibrated phylogenies of plant taxa, Bernal et al., 2019 proposed the hypothesis that the current occurrence of plants from estuarine or coastal zones in western Amazonia is the legacy of Miocene marine incursions in the region. Divergence times of populations of *Pachira aquatica* (Malvaceae) and *Manicaria saccifera* (Arecaceae) suggest that Amazonian and costal populations differentiated as a result of the complete retreat of the marine-influenced embayment. Appropiate testing of this hypothesis to reconstruct the extent of past estuarine environments requires further geogenomics studies, for which Bernal et al., 2019 provide a list of 28 candidate species from deltaic or estuarine environments in the region. Evidence for past estuarine connections from botanical data also include a study of a relict inland mangrove ecosystem (*Rhizophora mangle*) in the Yucatán peninsula (Aburto-Oropeza et al., 2021). Population genetic analyses from genotyping by sequencing data, together with floristics, sedimentological, paleontological evidence, and sea-level modeling, point to the mangrove relict reaching its current inland location in the Yucatán peninsula during last interglacial (ca. 120 K years).

FUTURE CONSIDERATIONS FOR PLANT GEOGENOMICS

Geogenomics research is a call to conduct research that specifically transfers information from biology to geology, making use of the wealth of information that can be extracted from large-scale genomic data. This sets it apart from other fields like comparative phylogeography, which in search of finding landscape and climatic features that drive shared genomic splits across co-distributed taxa (Edwards et al., 2022), could in fact inform geologic models without necessarily attempting to explicitly test models of landscape or climatic evolution. In addition to a paradigm shift to explicitly constrain geomorphic hypotheses, geogenomics research advocates for interdisciplinary networks. An example of this is the Baja GeoGenomics consortium (BGGc; https://bajageogenomics.org), a research group of geologists, conservation geneticists, and evolutionary biologists working with mammals, reptiles, and plants. The group aims to constrain how the physical landscape of the Baja California Peninsula has changed over the last 6 My, and its impact on the evolutionary trajectories of organisms in the region.

Given their limited dispersal, distribution patterns, and affinity to specific habitats, many plant groups are excellent systems for geogenomics research. Recent botanical investigations demonstrate that the analysis of data derived from plant DNA can improve our knowledge of paleotopography (Sanín et al., 2022), river system structure (Sander et al., 2018; Bedoya et al., 2021), habitat connections (Bernal et al., 2019; Aburto-Oropeza et al., 2021), or biome configuration through time (Särkinen et al., 2012). In times when advancing technologies allow the generation of massive amounts of genomic data, classical botanical knowledge is critical to identify good study systems to unveil long-standing questions on how the earth evolved with the use of plant DNA.

ACKNOWLEDGEMENTS

I am very thankful to several people from evolutionary biologists and botanists to paleontologists, whose contributions have helped shape my approach to questions lying at the intersection of botany and geology. Laura Lagomarsino, Carrie Tribble, Dick Olmstead, Luke Weaver, and an anonymous reviewer provided discussion and suggestions that helped improve this manuscript.

REFERENCES

Aburto-Oropeza, O., C. M. Burelo-Ramos, E. Ezcurra, P. Ezcurra, C. L. Henriquez, S. E. Vanderplank, and F. Zapata. 2021. Relict inland mangrove ecosystem reveals Last Interglacial sea levels. *Proceedings of the National Academy of Sciences* 118: e2024518118.

Adams, A. M., and R. R. Hudson. 2004. Maximum-Likelihood Estimation of Demographic Parameters Using the Frequency Spectrum of Unlinked Single-Nucleotide Polymorphisms. *Genetics* 168: 1699–1712.

Albert, J. S., N. R. Lovejoy, and W. G. R. Crampton. 2006. Miocene tectonism and the separation of cis- and trans-Andean river basins: Evidence from Neotropical fishes. *Journal of South American Earth Sciences* 21: 14–27.

Anderson, V. J., B. K. Horton, J. E. Saylor, A. Mora, E. Tesón, D. O. Breecker, and R. A. Ketcham. 2016. Andean topographic growth and basement uplift in southern Colombia: Implications for the evolution of the Magdalena, Orinoco, and Amazon river systems. *Geosphere* 12: 1235–1256.

Antonelli, A., J. A. A. Nylander, C. Persson, and I. Sanmartin. 2009. Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proceedings of the National Academy of Sciences* 106: 9749–9754.

Antonelli, A., and I. Sanmartín. 2011. Why are there so many plant species in the Neotropics? *TAXON* 60: 403–414.

Araya-Donoso, R., S. M. Baty, P. Alonso-Alonso, M. J. Sanín, B. T. Wilder, A. Munguia-Vega, and G. A. Dolby. 2022. Implications of barrier ephemerality in geogenomic research. *Journal of Biogeography* 49: 2050–2063.

Bacon, C. D., D. Silvestro, C. Jaramillo, B. T. Smith, P. Chakrabarty, and A. Antonelli. 2015. Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proceedings of the National Academy of Sciences* 112: 6110–6115.

Baker, P. A., S. C. Fritz, C. W. Dick, A. J. Eckert, B. K. Horton, S. Manzoni, C. C. Ribas, et al. 2014. The emerging field of geogenomics: Constraining geological problems with genetic data. *Earth-Science Reviews* 135: 38–47.

Bedoya, A. M., A. D. Leaché, and R. G. Olmstead. 2021. Andean uplift, drainage basin formation, and the evolution of plants living in fast-flowing aquatic ecosystems in northern South America. *New Phytologist*: nph.17649.

Bell, C. 2005. Phylogeny and biogeography of Valerianaceae (Dipsacales) with special reference to the South American valerians. *Organisms Diversity & Evolution* 5: 147–159.

Bernal, R., C. D. Bacon, H. Balslev, C. Hoorn, S. J. Bourlat, H. Tuomisto, S. Salamanca, et al. 2019. Could coastal plants in western Amazonia be relicts of past marine incursions? *Journal of Biogeography* 46: 1749–1759.

Blischak, P. D., J. Chifman, A. D. Wolfe, and L. S. Kubatko. 2018a. HyDe: A Python Package for Genome-Scale Hybridization Detection D. Posada [ed.],. *Systematic Biology* 67: 821–829.

Blischak, P. D., L. S. Kubatko, and A. D. Wolfe. 2018b. SNP genotyping and parameter estimation in polyploids using low-coverage sequencing data B. Berger [ed.],. *Bioinformatics* 34: 407–415.

Blischak, P. D., M. Sajan, M. S. Barker, and R. N. Gutenkunst. 2023. Demographic History Inference and the Polyploid Continuum S. Gravel [ed.],. *GENETICS*: iyad107.

Boschman, L. M. 2021. Andean mountain building since the Late Cretaceous: A paleoelevation reconstruction. *Earth-Science Reviews*: 103640.

Clague, D. A., and D. R. Sherrod. 2014. Growth and degradation of Hawaiian volcanoes: Chapter 3 in Characteristics of Hawaiian volcanoes. US Geological Survey.

Clark, A. G. 1997. Neutral behavior of shared polymorphism. *Proceedings of the National Academy of Sciences* 94: 7730–7734.

Clark, L. V., A. E. Lipka, and E. J. Sacks. 2019. polyRAD: Genotype Calling with Uncertainty from Sequencing Data in Polyploids and Diploids. *G3 Genes|Genomes|Genetics* 9: 663–673.

Crane, P. R., P. Herendeen, and E. M. Friis. 2004. Fossils and plant phylogeny. *American Journal of Botany* 91: 1683–1699.

Davis, C. C., C. D. Bell, S. Mathews, and M. J. Donoghue. 2002. Laurasian migration explains Gondwanan disjunctions: Evidence from Malpighiaceae. *Proceedings of the National Academy of Sciences* 99: 6833–6837.

Dewey, J. F., R. E. Holdsworth, and R. A. Strachan. 1998. Transpression and transtension zones. *Geological Society, London, Special Publications* 135: 1–14.

Dolby, G. A., S. E. K. Bennett, R. J. Dorsey, M. F. Stokes, B. R. Riddle, A. Lira-Noriega, A. Munguia-Vega, and B. T. Wilder. 2022. Integrating Earth–life systems: a geogenomic approach. *Trends in Ecology & Evolution* 37: 371–384.

Donoghue, M. J., D. A. R. Eaton, C. A. Maya-Lastra, M. J. Landis, P. W. Sweeney, M. E. Olson, N. I. Cacho, et al. 2022. Replicated radiation of a plant clade along a cloud forest archipelago. *Nature Ecology & Evolution* 6: 1318–1329.

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

Durand, E. Y., N. Patterson, D. Reich, and M. Slatkin. 2011. Testing for Ancient Admixture between Closely Related Populations. *Molecular Biology and Evolution* 28: 2239–2252.

Edwards, S. V., V. V. Robin, N. Ferrand, and C. Moritz. 2022. The Evolution of Comparative Phylogeography: Putting the Geography (and More) into Comparative Population Genomics R. Corbett-Detig [ed.],. *Genome Biology and Evolution* 14: evab176.

Farleigh, K., S. A. Vladimirova, C. Blair, J. T. Bracken, N. Koochekian, D. R. Schield, D. C. Card, et al. 2021. The effects of climate and demographic history in shaping genomic variation across populations of the Desert Horned Lizard (*Phrynosoma platyrhinos*). *Molecular Ecology* 30: 4481–4496.

Flantua, S. G. A., A. O'Dea, R. E. Onstein, C. Giraldo, and H. Hooghiemstra. 2019. The flickering connectivity system of the north Andean páramos. *Journal of Biogeography* 46: 1808–1825.

Frailey, C. D., E. L. Lavina, A. Rancy, and J. P. D. Souza Filho. 1988. A proposed Pleistocene/Holocene lake in the Amazon basin and its significance to Amazonian geology and biogeography. *Acta Amazonica* 18: 119–143.

Freyman, W. A., M. G. Johnson, and C. J. Rothfels. 2023. homologizer: Phylogenetic phasing of gene copies into polyploid subgenomes. *Methods in Ecology and Evolution* 14: 1230–1244.

Frost, L. A., S. M. Tyson, P. Lu-Irving, N. O'Leary, and R. G. Olmstead. 2017. Origins of North American arid-land Verbenaceae: More than one way to skin a cat. *American Journal of Botany* 104: 1708–1716.

Frost, L., D. A. Santamaría-Aguilar, D. Singletary, and L. P. Lagomarsino. 2022. Neotropical niche evolution of *Otoba* trees in the context of global biogeography of the nutmeg family. *Journal of Biogeography* 49: 156–170.

Gardner, E. M., M. G. Johnson, J. T. Pereira, A. S. A. Puad, D. Arifiani, Sahromi, N. J. Wickett, and N. J. C. Zerega. 2021. Paralogs and Off-Target Sequences Improve Phylogenetic Resolution in a Densely Sampled Study of the Breadfruit Genus (*Artocarpus* , Moraceae) M. Charleston [ed.],. *Systematic Biology* 70: 558–575.

Garrison, E., and G. Marth. 2012. Haplotype-based variant detection from short-read sequencing.

Garzione, C. N., G. D. Hoke, J. C. Libarkin, S. Withers, B. MacFadden, J. Eiler, P. Ghosh, and A. Mulch. 2008. Rise of the Andes. *Science* 320: 1304–1307.

Gentry, A. H. 1982. Neotropical Floristic Diversity: Phytogeographical Connections Between Central and South America, Pleistocene Climatic Fluctuations, or an Accident of the Andean Orogeny? *Annals of the Missouri Botanical Garden* 69: 557.

Gerard, D., L. F. V. Ferrão, A. A. F. Garcia, and M. Stephens. 2018. Genotyping Polyploids from Messy Sequencing Data. *Genetics* 210: 789–807.

Gianni, G. M., C. Navarrete, A. Echaurren, M. Díaz, K. L. Butler, B. K. Horton, A. Encinas, and A. Folguera. 2020. Northward propagation of Andean genesis: Insights from Early Cretaceous synorogenic deposits in the Aysén-Río Mayo basin. *Gondwana Research* 77: 238–259.

Givnish, T. J., M. H. J. Barfuss, B. Van Ee, R. Riina, K. Schulte, R. Horres, P. A. Gonsiska, et al. 2011. Phylogeny, adaptive radiation, and historical biogeography in Bromeliaceae: Insights from an eight-locus plastid phylogeny. *American Journal of Botany* 98: 872–895.

Givnish, T. J., K. C. Millam, A. R. Mast, T. B. Paterson, T. J. Theim, A. L. Hipp, J. M. Henss, et al. 2009. Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). *Proceedings of the Royal Society B: Biological Sciences* 276: 407–416.

Givnish, T. J., D. Spalink, M. Ames, S. P. Lyon, S. J. Hunter, A. Zuluaga, A. Doucette, et al. 2016. Orchid historical biogeography, diversification, Antarctica and the paradox of orchid dispersal. *Journal of Biogeography* 43: 1905–1916.

Green, R. E., J. Krause, A. W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, et al. 2010. A Draft Sequence of the Neandertal Genome. *Science* 328: 710–722.

Gregory-Woodzicki, K. M. 2000. Uplift history of the Central and Northern Andes: A Review. 112: 1091–1105.

Gutenkunst, R. N., R. D. Hernandez, S. H. Williamson, and C. D. Bustamante. 2009. Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data G. McVean [ed.],. *PLoS Genetics* 5: e1000695.

Hoorn, C. 2006. Mangrove Forests and Marine Incursions in Neogene Amazonia (Lower Apaporis River, Colombia). *PALAIOS* 21: 197–209.

Hoorn, C. 1993. Marine incursions and the influence of Andean tectonics on the Miocene depositional history of northwestern Amazonia: results of a palynostratigraphic study. *Palaeogeography, Palaeoclimatology, Palaeoecology* 105: 267–309.

Hoorn, C., A. Perrigo, and A. Antonelli. 2018. Mountains, climate and biodiversity. Wiley, Hoboken.

Hoorn, C., F. P. Wesselingh, H. ter Steege, M. A. Bermudez, A. Mora, J. Sevink, I. Sanmartin, et al. 2010. Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity. *Science* 330: 927–931.

Horton, B. K. 2018. Sedimentary record of Andean mountain building. *Earth-Science Reviews* 178: 279–309.

Hughes, C. E., R. T. Pennington, and A. Antonelli. 2013. Neotropical Plant Evolution: Assembling the Big Picture: Neotropical Plant Evolution. *Botanical Journal of the Linnean Society* 171: 1–18.

Hughes, C., and R. Eastwood. 2006. Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences* 103: 10334–10339.

Humboldt, A. V., and A. Bonpland. 1807. Essai sur la géographie des plantes, accompagné d'un tableau physique des régions équinoxiales. Levrault & Schoell, Paris.

Jaramillo, C. 2023. The evolution of extant South American tropical biomes. *New Phytologist* 239: 477–493.

Jaramillo, C., I. Romero, C. D'Apolito, G. Bayona, E. Duarte, S. Louwye, J. Escobar, et al. 2017. Miocene flooding events of western Amazonia. *Science Advances* 3: e1601693.

Ji, J., D. J. Jackson, A. D. Leaché, and Z. Yang. 2023. Power of Bayesian and Heuristic Tests to Detect Cross-Species Introgression with Reference to Gene Flow in the *Tamias quadrivittatus* Group of North American Chipmunks D. R. Thompson [ed.],. *Systematic Biology* 72: 446–465.

Johnson, M. G., E. M. Gardner, Y. Liu, R. Medina, B. Goffinet, A. J. Shaw, N. J. C. Zerega, and N. J. Wickett. 2016. HybPiper: Extracting Coding Sequence and Introns for Phylogenetics from High-Throughput Sequencing Reads Using Target Enrichment. *Applications in Plant Sciences* 4: 1600016.

Kolář, F. 2021. Tracing evolutionary history and admixture in mixed-ploidy systems. *Molecular Ecology Resources* 21: 1413–1415.

Lagomarsino, L. P., F. L. Condamine, A. Antonelli, A. Mulch, and C. C. Davis. 2016. The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytologist* 210: 1430–1442.

Landis, M., E. J. Edwards, and M. J. Donoghue. 2021. Modeling Phylogenetic Biome Shifts on a Planet with a Past M. Holder [ed.],. *Systematic Biology* 70: 86–107.

Landis, M. J. 2017. Biogeographic Dating of Speciation Times Using Paleogeographically Informed Processes. *Systematic Biology*: syw040.

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

Latrubesse, E. M., M. Cozzuol, S. A. F. da Silva-Caminha, C. A. Rigsby, M. L. Absy, and C. Jaramillo. 2010. The Late Miocene paleogeography of the Amazon Basin and the evolution of the Amazon River system. *Earth-Science Reviews* 99: 99–124.

Latrubesse, E. M., S. A. F. Da Silva, M. Cozzuol, and M. L. Absy. 2007. Late Miocene continental sedimentation in southwestern Amazonia and its regional significance: Biotic and geological evidence. *Journal of South American Earth Sciences* 23: 61–80.

Lawson, D. J., L. van Dorp, and D. Falush. 2018. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nature Communications* 9: 3258.

Li, H., and R. Durbin. 2011. Inference of human population history from individual whole-genome sequences. *Nature* 475: 493–496.

Lim, J. Y., and C. R. Marshall. 2017. The true tempo of evolutionary radiation and decline revealed on the Hawaiian archipelago. *Nature* 543: 710–713.

Linhares, A. P., V. D. C. D. S. Gaia, and M. I. F. Ramos. 2017. The significance of marine microfossils for paleoenvironmental reconstruction of the Solimões Formation (Miocene), western Amazonia, Brazil. *Journal of South American Earth Sciences* 79: 57–66.

Loog, L. 2021. Sometimes hidden but always there: the assumptions underlying genetic inference of demographic histories. *Philosophical Transactions of the Royal Society B: Biological Sciences* 376: 20190719.

Mandel, J. R., R. B. Dikow, C. M. Siniscalchi, R. Thapa, L. E. Watson, and V. A. Funk. 2019. A fully resolved backbone phylogeny reveals numerous dispersals and explosive diversifications throughout the history of Asteraceae. *Proceedings of the National Academy of Sciences* 116: 14083–14088.

Mendez-Reneau, J., J. Gordon Burleigh, and E. M. Sigel. 2023. Target Capture Methods Offer Insight into the Evolution of Rapidly Diverged Taxa and Resolve Allopolyploid Homeologs in the Fern Genus *Polypodium* s.s. *Systematic Botany* 48: 96–109.

Montes, C., A. F. Rodriguez-Corcho, G. Bayona, N. Hoyos, S. Zapata, and A. Cardona. 2019. Continental margin response to multiple arc-continent collisions: The northern Andes-Caribbean margin. *Earth-Science Reviews* 198: 102903.

Montes, C., C. A. Silva, G. A. Bayona, R. Villamil, E. Stiles, A. F. Rodriguez-Corcho, A. Beltran-Triviño, et al. 2021. A Middle to Late Miocene Trans-Andean Portal: Geologic Record in the Tatacoa Desert. *Frontiers in Earth Science* 8: 587022.

Mora, A., P. Baby, M. Roddaz, M. Parra, S. Brusset, W. Hermoza, and N. Espurt. 2010. Tectonic history of the Andes and sub-Andean zones: implications for the development of the

Amazon drainage basin. Amazonia, Landscape and Species Evolution, 38–60. Wiley, Oxford.

Mora, A., M. Parra, M. R. Strecker, E. R. Sobel, H. Hooghiemstra, V. Torres, and J. V. Jaramillo. 2008. Climatic forcing of asymmetric orogenic evolution in the Eastern Cordillera of Colombia. *Geological Society of America Bulletin* 120: 930–949.

Morales-Briones, D. F., B. Gehrke, C.-H. Huang, A. Liston, H. Ma, H. E. Marx, D. C. Tank, and Y. Yang. 2022. Analysis of Paralogs in Target Enrichment Data Pinpoints Multiple Ancient Polyploidy Events in *Alchemilla* s.l. (Rosaceae) J. Mandel [ed.],. *Systematic Biology* 71: 190–207.

Mora-Páez, H., J. N. Kellogg, J. T. Freymueller, D. Mencin, R. M. S. Fernandes, H. Diederix, P. LaFemina, et al. 2019. Crustal deformation in the northern Andes – A new GPS velocity field. *Journal of South American Earth Sciences* 89: 76–91.

Nauheimer, L., N. Weigner, E. Joyce, D. Crayn, C. Clarke, and K. Nargar. 2021. HybPhaser: A workflow for the detection and phasing of hybrids in target capture data sets. *Applications in Plant Sciences* 9: aps3.11441.

Nazareno, A. G., C. W. Dick, and L. G. Lohmann. 2017. Wide but not impermeable: Testing the riverine barrier hypothesis for an Amazonian plant species. *Molecular Ecology* 26: 3636–3648.

Nevado, B., N. Contreras-Ortiz, C. Hughes, and D. A. Filatov. 2018. Pleistocene glacial cycles drive isolation, gene flow and speciation in the high-elevation Andes. *New Phytologist* 219: 779–793.

Noriega-Londoño, S., S. A. Restrepo-Moreno, C. Vinasco, M. A. Bermúdez, and K. Min. 2020. Thermochronologic and geomorphometric constraints on the Cenozoic landscape evolution of the Northern Andes: Northwestern Central Cordillera, Colombia. *Geomorphology* 351: 106890.

Pardi, F., and C. Scornavacca. 2015. Reconstructible Phylogenetic Networks: Do Not Distinguish the Indistinguishable B. R. Holland [ed.],. *PLOS Computational Biology* 11: e1004135.

Picq, S., F. Alda, R. Krahe, and E. Bermingham. 2014. Miocene and Pliocene colonization of the Central American Isthmus by the weakly electric fish *Brachyhypopomus occidentalis* (Hypopomidae, Gymnotiformes) M. Ebach [ed.],. *Journal of Biogeography* 41: 1520–1532.

Rahbek, C., M. K. Borregaard, A. Antonelli, R. K. Colwell, B. G. Holt, D. Nogues-Bravo, C. M. Ø. Rasmussen, et al. 2019. Building mountain biodiversity: Geological and evolutionary processes. *Science* 365: 1114–1119.

Räsänen, M. E., A. M. Linna, J. C. R. Santos, and F. R. Negri. 1995. Late Miocene Tidal Deposits in the Amazonian Foreland Basin. *Science* 269: 386–390.

Rodriguez-Muñoz, E., C. Montes, F. J. M. Rojas-Runjaic, and A. J. Crawford. 2022. Synthesis of geological data and comparative phylogeography of lowland tetrapods suggests recent dispersal through lowland portals crossing the Eastern Andean Cordillera. *PeerJ* 10: e13186.

Roquet, C., F. C. Boucher, W. Thuiller, and S. Lavergne. 2013. Replicated radiations of the alpine genus *Androsace* (Primulaceae) driven by range expansion and convergent key innovations A. Phillimore [ed.],. *Journal of Biogeography* 40: 1874–1886.

Rose, J. P., C.-L. Xiang, K. J. Sytsma, and B. T. Drew. 2022. A timeframe for mint evolution: towards a better understanding of trait evolution and historical biogeography in Lamiaceae. *Botanical Journal of the Linnean Society* 200: 15–38.

Roux, C., and J. R. Pannell. 2015. Inferring the mode of origin of polyploid species from next-generation sequence data. *Molecular Ecology* 24: 1047–1059.

Roux, C., X. Vekemans, and J. Pannell. 2023. Inferring the Demographic History and Inheritance Mode of Tetraploid Species Using ABC. *In* Y. Van De Peer [ed.], Polyploidy, Methods in Molecular Biology, 325–348. Springer US, New York, NY.

Ruokolainen, K., G. M. Moulatlet, G. Zuquim, C. Hoorn, and H. Tuomisto. 2019. Geologically recent rearrangements in central Amazonia river network and their importance for the riverine barrier hypothesis. 11: e45046.

Sander, N. L., F. Pérez-Zavala, C. J. Da Silva, J. C. Arruda, M. T. Pulido, M. A. A. Barelli, A. B. Rossi, et al. 2018. Rivers shape population genetic structure in *Mauritia flexuosa* (Arecaceae). *Ecology and Evolution* 8: 6589–6598.

Sanín, M. J., F. G. Mejía-Franco, M. Paris, W. A. Valencia-Montoya, N. Salamin, M. Kessler, I. Olivares, et al. 2022. Geogenomics of montane palms points to MIOCENE–PLIOCENE Andean segmentation related to strike-slip tectonics. *Journal of Biogeography* 49: 1711–1725.

Särkinen, T., R. T. Pennington, M. Lavin, M. F. Simon, and C. E. Hughes. 2012. Evolutionary islands in the Andes: persistence and isolation explain high endemism in Andean dry tropical forests: Evolutionary islands in the Andes. *Journal of Biogeography* 39: 884–900.

Schenk, J. J. 2016. Consequences of Secondary Calibrations on Divergence Time Estimates A. Janke [ed.],. *PLOS ONE* 11: e0148228.

Schiffels, S., and R. Durbin. 2014. Inferring human population size and separation history from multiple genome sequences. *Nature Genetics* 46: 919–925.

Schoene, B., D. J. Condon, L. Morgan, and N. McLean. 2013. Precision and Accuracy in Geochronology. *Elements* 9: 19–24.

Serang, O., M. Mollinari, and A. A. F. Garcia. 2012. Efficient Exact Maximum a Posteriori Computation for Bayesian SNP Genotyping in Polyploids F. Rapallo [ed.],. *PLoS ONE* 7: e30906.

Shastry, V., P. E. Adams, D. Lindtke, E. G. Mandeville, T. L. Parchman, Z. Gompert, and C. A. Buerkle. 2021. Model-based genotype and ancestry estimation for potential hybrids with mixed-ploidy. *Molecular Ecology Resources* 21: 1434–1451.

Sklenář, P., E. Dušková, and H. Balslev. 2011. Tropical and Temperate: Evolutionary History of Páramo Flora. *The Botanical Review* 77: 71–108.

Smith, B. T., J. E. McCormack, A. M. Cuervo, Michael. J. Hickerson, A. Aleixo, C. D. Cadena, J. Pérez-Emán, et al. 2014. The drivers of tropical speciation. *Nature* 515: 406–409.

Solís-Lemus, C., and C. Ané. 2016. Inferring Phylogenetic Networks with Maximum Pseudolikelihood under Incomplete Lineage Sorting S. Edwards [ed.],. *PLOS Genetics* 12: e1005896.

Solís-Lemus, C., P. Bastide, and C. Ané. 2017. PhyloNetworks: A Package for Phylogenetic Networks. *Molecular Biology and Evolution* 34: 3292–3298.

Stebbins, G. L. 1950. Variation and Evolution in Plants. Columbia University Press.

Stift, M., F. Kolář, and P. G. Meirmans. 2019. Structure is more robust than other clustering methods in simulated mixed-ploidy populations. *Heredity* 123: 429–441.

Sundaram, M., M. J. Donoghue, A. Farjon, D. Filer, S. Mathews, W. Jetz, and A. B. Leslie. 2019. Accumulation over evolutionary time as a major cause of biodiversity hotspots in conifers. *Proceedings of the Royal Society B: Biological Sciences* 286: 20191887.

Tagliacollo, V. A., F. F. Roxo, S. M. Duke-Sylvester, C. Oliveira, and J. S. Albert. 2015. Biogeographical signature of river capture events in Amazonian lowlands. *Journal of Biogeography* 42: 2349–2362.

Terhorst, J., and Y. S. Song. 2015. Fundamental limits on the accuracy of demographic inference based on the sample frequency spectrum. *Proceedings of the National Academy of Sciences* 112: 7677–7682.

Torres Jimenez, M. F., G. N. Stone, A. Sanchez, and J. E. Richardson. 2021. Comparative phylogeography of an ant-plant mutualism: An encounter in the Andes. *Global and Planetary Change* 205: 103598.

Velandia, F., J. Acosta, R. Terraza, and H. Villegas. 2005. The current tectonic motion of the Northern Andes along the Algeciras Fault System in SW Colombia. *Tectonophysics* 399: 313–329.

Veloza, G., R. Styron, M. Taylor, and A. Mora. 2012. Open-source archive of active faults for northwest South America. *GSA Today* 22: 4–10.

Vonhof, H. B., F. P. Wesselingh, R. J. G. Kaandorp, G. R. Davies, J. E. Van Hinte, J. Guerrero, M. Räsänen, et al. 2003. Paleogeography of Miocene Western Amazonia: Isotopic composition of molluscan shells constrains the influence of marine incursions. *Geological Society of America Bulletin* 115: 983–993.

Webb, S. D. 1995. Biological Implications of the Middle Miocene Amazon Seaway. *Science* 269: 361–362.

Wesselingh, F. P., M. E. Räsänen, G. Irion, H. B. Vonhof, R. Kaandorp, W. Renema, L. Romero Pittman, and M. Gingras. 2002. Lake Pebas: a palaeoecological reconstruction of a Miocene, lon-lived lake complex in western Amazonia. 1: 35–81.

Yang, Y., and S. A. Smith. 2014. Orthology Inference in Nonmodel Organisms Using Transcriptomes and Low-Coverage Genomes: Improving Accuracy and Matrix Occupancy for Phylogenomics. *Molecular Biology and Evolution* 31: 3081–3092.

Yu, Y., J. Dong, K. J. Liu, and L. Nakhleh. 2014. Maximum likelihood inference of reticulate evolutionary histories. *Proceedings of the National Academy of Sciences* 111: 16448–16453.

Zapata, S., L. Calderon-Diaz, C. Jaramillo, F. Oboh-Ikuenobe, J. C. Piedrahita, M. Rodríguez-Cuevas, A. Cardona, et al. 2023. Drainage and sedimentary response of the Northern Andes and the Pebas system to Miocene strike-slip tectonics: A source to sink study of the Magdalena Basin. *Basin Research*: bre.12769.