Trends in
Ecology & Evolution

CellPress

Review

Integrating Earth–life systems: a geogenomic approach

Greer A. Dolby, [1,2,3,*] Scott E.K. Bennett, [3,4,*] Rebecca J. Dorsey, [3,5] Maya F. Stokes, [3,6] Brett R. Riddle, [3,7] Andrés Lira-Noriega, [3,8] Adrian Munguia-Vega, [3,9] and Benjamin T. Wilder [3,9,*]

For centuries, scientists have recognized and worked to understand how Earth's mutable landscape and climate shape the distribution and evolution of species. Here, we describe the emerging field of geogenomics, which uses the reciprocal and deep integration of geologic, climatic, and population genomic data to define and test cause–effect relationships between Earth and life at intermediate spatial and temporal scales (i.e., the mesoscale). Technological advances now power the detailed reconstruction of landscape and evolutionary histories, but transdisciplinary collaborations and new quantitative tools are needed to better integrate Earth–life data. Geogenomics can help build a more unified theory and characterize the boundary conditions under which geologic and climatic processes generate new biodiversity, how species' responses differ, and why.

Highlights

Considering Earth's complex and co-occurring geologic and climatic processes is important for understanding species evolution.

Use of genome-scale sequence data helps identify which geologic and climatic forces shaped adaptation and divergence of populations and sister species.

The emerging field of geogenomics emphasizes independent, testing of non-mutually exclusive hypotheses using deep, reciprocal integration of Earth and evolutionary sciences.

Geogenomics offers a new framework for understanding the boundary conditions under which geologic and climatic processes shape the evolution of species.

A geogenomics approach can reveal the biological traits that make species more or less responsive to landscape and climate change.

Reconnecting the Earth–life sciences

> ...A great light has been thrown upon [species geographical distributions] by geological investigations, which have shown that the present state of the earth, and the organisms now inhabiting it, are but the last stage of a long and uninterrupted series of changes [and to] account for its present condition without any reference to those changes [...] must lead to very imperfect and erroneous conclusions. Alfred Russel Wallace

The insights that 19th century scientists Humboldt, Lyell, Wallace, and Darwin contributed to evolutionary biology were founded on the perspective that the physical properties of Earth and the biological patterns of plants and animals are linked. In the following century, these disciplines specialized into separate branches that, along with technological advances, yielded revolutions such as plate tectonics theory and whole-genome sequencing. While that specialization spurred countless important discoveries, it has also impeded the transmission of knowledge between the two fields [1].

Future breakthroughs toward understanding the 'unity of nature' [2] will require weaving together independent knowledge domains to reconnect the Earth and life sciences. Such a reunification today comes powered by 21st century technologies, quantitative frameworks, and disciplinary insights that can together answer long-standing questions in new ways. We are now poised to re-integrate these disciplines with a mechanism-focused perspective, inspired by the questions of past generations and the technology revolution of the present.

Earth–life research at the mesoscale

Geologic, climatic, and biologic processes modulate Earth's surface and atmosphere over many temporal and spatial scales. However, it is Earth's dynamic nature over thousand- to million-year

[1]School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA
[2]Center for Mechanisms of Evolution, Biodesign Center, Arizona State University, Tempe, AZ 85287, USA
[3]Baja GeoGenomics Consortium, www.bajageogenomics.org
[4]U.S. Geological Survey, Geology, Minerals, Energy, and Geophysics Science Center, Portland, OR 97201, USA
[5]Department of Earth Sciences, University of Oregon, Eugene, OR, 97403, USA
[6]Yale Institute for Biospheric Studies, Yale University, New Haven, CT 06511, USA
[7]School of Life Sciences, University of Nevada Las Vegas, Las Vegas, NV 89154, USA
[8]CONACyT Research Fellow, Red de Estudios Moleculares Avanzados, Instituto de Ecología, A.C., Xalapa, Veracruz 91073, Mexico

timescales and at regional spatial scales that fundamentally shape the incipient speciation and distribution of species. Several disciplines within biology seek to understand how the physical and climatic **landscape** (see Glossary) shapes patterns of species divergence and diversity. At the highest level of organization (**macroscale**), **phylogenetic** and **macroecology/biodiversity informatics** studies provide complementary information about the distribution and diversity patterns of species over continental to global scales [3,4] (Figure 1). For example, species richness data reveal what geographic settings associate with high richness and among which taxonomic groups [5]. These results motivate hypotheses of what geologic or climatic features have led to observed richness patterns [6]. However, they cannot determine whether the relationship is due to the accumulation or loss of biodiversity through dispersal or low extinction rates versus due to speciation *in situ*. Phylogenies offer complementary insights to reveal whether lineage diversification occurred during a geologic or climatic change and whether co-distributed species diverged simultaneously [7]. Phylogenetic studies, however, have greater uncertainty regarding causes of divergence — they are better suited to confirm it as a plausible explanation if the timing of the two aligns. There are two reasons for this.

First, many settings are complicated by **pseudocongruence**, wherein more than one aspect of the physical landscape has changed over the evolutionary period of interest such that cause–effect relationships cannot be deduced from timing information alone. Second, even if an evolutionary pattern can be confidently assigned to a geologic or climatic change, it does not reveal what characteristic of that physical change was key for diversification or whether different species responded to the same characteristic [8,9]. Answering cause–effect questions about speciation requires another class of data. We propose that formalizing and testing cause–effect relationships about speciation is best achieved at the **mesoscale**, which we loosely define as the temporal and spatial scale at which population divergence and incipient speciation occur.

Within ecology, the mesoscale composes an uncertain mix of processes from higher and lower scales [10]. However, within evolutionary biology, the use of population genomic data at the mesoscale matches the spatiotemporal scale at which features on the landscape change (e.g., a mountain range, a drainage network). New geological methods allow us to characterize changes critical to Earth's climate and surface at this scale. Likewise, high-throughput sequencing allows assessment of detailed evolutionary responses within actively diverging populations at this scale. Thus, the spatiotemporal alignment of Earth and life datasets makes the mesoscale key to characterizing mechanisms by which geologic and climatic processes control speciation specifically (Figure 1). While this scale is similar to **phylogeography** or landscape genetics, we detail here a set of integrated methods, data, and questions that reach beyond these fields.

Working at the mesoscale allows the identification of cause–effect relationships between Earth processes and biological patterns because we can resolve both the stage and driver of divergence across diverging populations and derive **mechanistic** knowledge about such relationships. A main challenge is the deep communication required across fields with different nomenclature and historical norms. In addition, the financial and personnel resources required for such genomic studies are currently high (though decreasing) for even a limited number of taxa. However, mechanistic knowledge should explain patterns in other geographic and taxonomic settings broadly, yielding transformative insights.

Defining and testing geogenomic hypotheses

In this article, we extend the definition formalized in [11] and define **geogenomics** as the reciprocal deep integration of geologic, climatic, and genomic data to test independent hypotheses about the co-evolution of life and its landscapes over thousand- to million year timescales. We

[9]Desert Laboratory on Tumamoc Hill, University of Arizona, Tucson, AZ 85745, USA

*Correspondence: gadolby@asu.edu (G.A. Dolby), sekbennett@usgs.gov (S.E.K. Bennett), and bwilder@email.arizona.edu (B.T. Wilder).

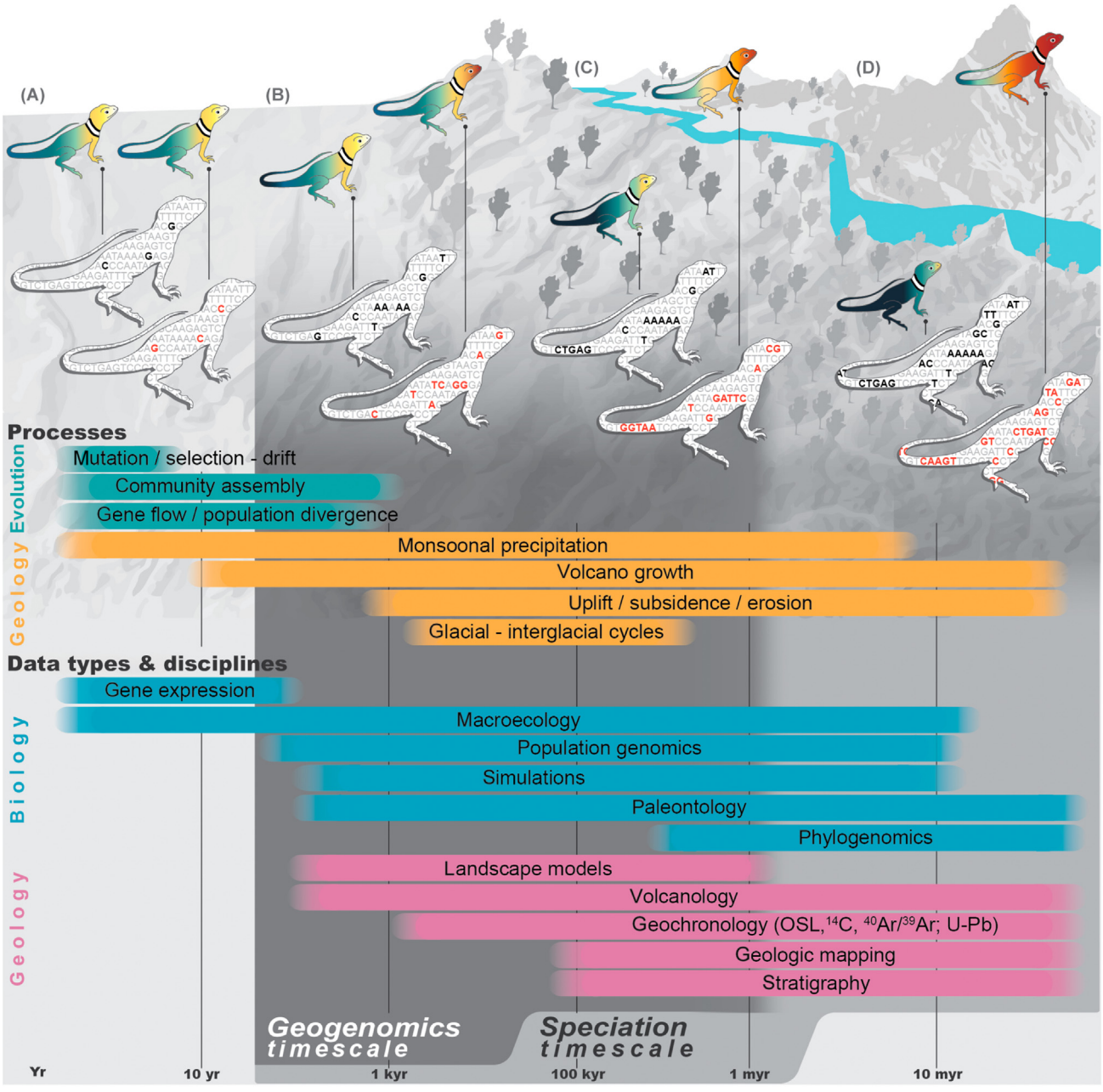

Figure 1. Concept diagram showing timescales of common biological and geological processes. Also shown are disciplines and data types relevant to speciation and geogenomic research. Time on the x axis is log scale and boundaries are approximations estimated to reflect the duration over which a process or data type is commonly studied. The amount of genetic divergence is illustrated by color and number of basepair mutations in the lizard silhouettes and increases to the right from steps A to D. The geogenomic approach outlined here targets the incipient portion of the speciation continuum (dark gray, bottom). Abbreviation: OSL, optically stimulated luminescence.

focus this article on integrating these data types to understand the external, abiotic drivers of incipient speciation using population genomic data. Our approach is complementary to the definition of geogenomics proposed by Baker, Fritz *et al.* [11] as mechanistic knowledge linking abiotic

processes to biotic diversification can ultimately better guide the use of genomic data to inform geologic and climatic hypotheses.

The first consideration of geogenomic studies is to determine the primary ways in which the physical setting (i.e., landscape) has changed over evolutionary timescales geologically and climatically (Box 1) [12,13]. This requires iteration between the Earth and life sciences. Ideally, the magnitude and geographic location of genetic divergence sets the spatiotemporal scale for Earth scientists to focus inquiries into past geologic and climatic history of a region. Characterization of these changes can reveal evolutionary hypotheses that are most plausible for biologists to test. This requires a close working relationship between disciplines to determine what biological patterns can be expected from geologic and climatic histories. For example, in order to test a causal hypothesis between a pattern of lineage divergence and landscape history, biologists may need to understand aspects of a river's history, such as its size, persistence, or erosional power during glaciations [14], or how high topographic features such as fault-related basins and ranges were when they formed and how that height changed over time. Such details are critical for biologists to properly choose data types, focal species, sampling areas and analyses to test evolutionary hypotheses. Similarly, geologists may need genomic expertise to interpret evolutionary patterns to properly inform geologic and climatic hypotheses. However, the details necessary to make these decisions are often insufficient in the literature and instead require conversations between scientists of different fields.

Reconstructing changes in Earth's surface

Earth's landscape plays a critical role in the distribution and evolution of species [15–22]. Geologic data provide fundamental constraints on the timing, magnitude, frequency, and duration of physical processes that reshape that landscape and landscape changes that impact speciation can be

Box 1. Baja California Peninsula: a geogenomic approach

The Baja California peninsula (BCP) hosts a complex geologic and climatic history and elevated levels of endemism, making it an ideal setting to test geogenomic hypotheses. BCP history includes tectonic rifting, subsidence and uplift, changing shorelines, glacial cycles, and a seasonal rainfall gradient superimposed along 10° of latitude from mid-latitude deserts to the tropics (Figure I). The BCP hosts a strong pattern of intraspecific north-south genetic divergence documented in over 60 disparate species [13]. Studies attributed this cryptic vicariance to a transpeninsular seaway that formed a temporary barrier to gene flow [102]. However, no geologic evidence of a seaway has been documented, and studies suggest that the BCP has been at its current elevation since ~3 Ma [103], challenging the seaway hypothesis and leaving the causal mechanism(s) unknown. Accordingly, a geogenomic approach focused on the BCP can evaluate which combination of geologic-climatic mechanisms drove evolutionary divergence.

Current research by the Baja GeoGenomics (BGG) consortium is testing three independent hypotheses on six exemplar desert taxa whose ranges span the N–S gradient to elucidate the drivers of the observed genetic divergence (Figure I). Each hypothesis is tested individually, allowing for the possibility of multiple, synergistic causal factors:

(i) Isolation by a seaway. Geological mapping and analysis of the central BCP will reconstruct the vertical history of the land and the evolving BCP shoreline to evaluate whether a transpeninsular seaway or a narrow isthmus existed in the past several million years. Genomic tests of whether N–S populations of species diverged at the same time will be used to evaluate whether a seaway, if supported geologically, could explain the present-day divergence patterns.

(ii) Isolation in glacial refugia. Whole genome analysis of N–S populations combined with ecological niche modeling will test whether habitat shifts during Pleistocene glacial–interglacial cycles isolated populations in fragmented refugia. Population genomic analysis will reveal signatures of range expansion or demographic history consistent with refugia and test expectations that glacial refugia identified by niche models are areas of high genetic diversity and form distinct groups.

(iii) Temporal isolation via climate dipole. Analysis of seasonal gene expression (transcriptomics) will reveal whether N–S populations diverged due to differences in rainfall timing (winter versus summer) driven by the North American monsoon [104]. Genome-wide tests for differential adaptation and niche divergence between N–S populations will evaluate whether divergence is additionally due to differential ecological adaptation.

Glossary

Biodiversity informatics: use of informatics and computational techniques to understanding biodiversity patterns.
Bottleneck: a strong decrease in the size of a biological population.
Climate dipole: contrasting pattern of regional climate conditions that appear at two proximate geographic locations at the same time.
Detrital: minerals or particles of rock derived from the mechanical breakdown of preexisting rocks by weathering and erosion.
Differential adaptation: process by which two populations or species adapt to different conditions, particularly environmental or ecological.
Extension: the tectonic process of stretching Earth's crust or lithosphere.
Fossil assemblage: a group of associated animal fossils found together in a sedimentary formation.
Functional elements: parts of the genome that play a role in biological structure or function, particularly parts of genes (e.g., exons) that produce proteins when expressed.
Gene ontology: knowledge (hosted in public databases) that describes the biological/chemical/molecular function of genes for all organisms.
General circulation models: a type of global 3D climate model used to predict weather and climate.
Genomic islands of differentiation: physical regions of the genome that are particularly different between study populations or species.
Geochemistry: the field of study concerned with the chemical composition of minerals, rocks, and geologic formations.
Geochronology: the field of study concerned with the dating of rocks, stratigraphic formations, and geological events.
Geogenomics: the study of how Earth and life coevolve over intermediate spatiotemporal scales through the deep integration of geologic, climatic, and genomic data to test independent hypotheses.
Glacio-eustatic sea-level change: change in sea level due to the uptake or release of water from glaciers and polar ice.
Landscape: the physical land–sea–atmosphere interface that makes up Earth's critical zone, including climate.

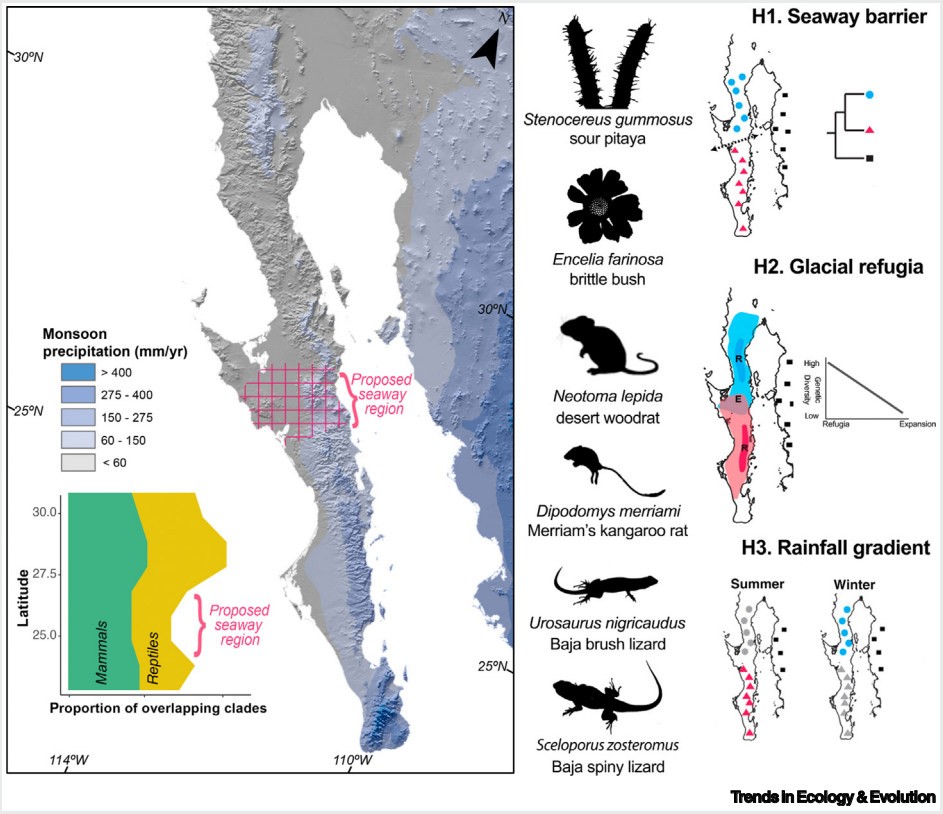

Figure I. (Left) Map of the Baja California Peninsula showing summer (July–September) monsoon precipitation, areas of high genetic diversity shown as proportion of overlapping genetic clades (bottom left inset), and purported Plio–Pleistocene seaways in low passes (pink crosshatched region). (Right) Illustrations of the six study species used to examine the three hypotheses along the Baja California peninsula by the Baja GeoGenomics (BGG) consortium.

quantified with the precision required to test evolutionary hypotheses. Important processes of this type include formation and destruction of topographic features such as mountain ranges, valleys, continental shelves, and ocean basins, carving of deep canyons, development of river drainages, and isolation of landmasses through the formation of islands or peninsulas. Genomic data can then test whether evolutionary consequences of species differ by dispersal ability or habitat type in these settings.

Technological advances allow Earth scientists to characterize these physical processes with increasingly high precision (Table 1). Methods in **geochronology** integrated with geologic mapping and field observations allow geologists to quantify the timing, duration, and rates of magma crystallization or a volcanic eruption using Ar/Ar geochronology, deposition of sediments by dating **detrital** zircon or sanidine minerals, landscape erosion and incision with **surface exposure dating**, mountain uplift or exhumation using low-temperature **thermochronology**, and flooding of seaways by identifying **fossil assemblages** in marine strata. Advances in process sedimentology, process volcanology, sequence stratigraphy, and **geochemistry** are used to reconstruct past environments and plate tectonic settings. High-resolution digital

Macroecology: the study of biological diversity using species occurrence data, typically on large spatial scales.
Macroscale: the largest spatial and/or temporal scale of study (continental-to-global, multi-million-year).
Mechanistic: knowledge regarding the elements and structures of a system or process that can explain how a phenomenon comes to be and which often guides theory of how/why the phenomenon exists.
Mesoscale: an intermediate (regional) spatial scale and thousand-to-million-year time scale of study on which incipient speciation is thought to occur.
Mutation burden: the accumulation of detrimental alleles within an individual that can affect that organism individual's fitness.
Neutral elements: parts of the genome that are not thought to be under direct natural selection or functional constraint (e.g., non-genic elements, repeat regions).
Orographic: the effects of mountains forcing moist air to rise.
Paleoaltimetry: the study of the ancient elevations of Earth's surface.
Phylogenetics: the study or use of tree networks based on molecular data that depict the relationships between organisms that share a common ancestor.
Phylogeography: a subdiscipline of biogeography that studies the geographic distribution of genealogical populations and species.
Pseudocongruence: a pattern that could be caused by more than one process often because those processes are nested in time and/or space and/or have similar effects.
Selective sweep: the genetic process in which a beneficial genetic variant rises to very high frequency within a population, often over a short time.
Subsidence: the lowering of Earth's crust and/or surface, for example due to extensional tectonics and sediment compaction; often form lowlands where sediments/water accumulate.
Surface exposure dating: a collection of geochronological techniques used to estimate the length of time a rock has been exposed at or near Earth's surface.
Temporal isolation: a phenomenon whereby biological populations do not interact/reproduce because they are active at different times instead of being physically isolated.

elevation data [e.g., Shuttle Radar Topography Mission (SRTM), lidar] permit rapid remote characterization of Earth's physiography and to interpret processes such as incision, uplift, and deposition that modify the landscape. Limits to our ability to reconstruct paleotopography are typically related to the spatiotemporal scale of changes studied and uncertainties of analytical techniques (Table 1). Other limitations relate to the spatially fragmented and temporally incomplete nature of the geologic record, which can reduce its interpretability.

By integrating geologic datasets, the past motions, elevations, and locations of mountain ranges or tectonic plates can be combined to reconstruct the paleolandscapes on which species evolved [23,24]. For example, integration of geologic mapping, cooling ages of exhumed rocks, and **paleoaltimetry** estimates based on leaf-wax lipids and volcanic glass isotopes show that the Andes Mountains in South America underwent uplift at spatially variable rates from ~25 to 10 million years ago (Ma) [25,26]. Similar studies show that the modern Basin and Range province in western North America consists of fragments of a formerly contiguous elevated plateau (Nevadaplano) that existed ~41–23 Ma, at elevations 1–2 km higher than present-day mountain ranges and was subsequently destroyed by tectonic **extension** and crustal **subsidence** [23,27]. Using such results, it should be possible to assess how subsidence rates or landscape fragmentation changed the niche space available to organisms, or selection pressure on altitude-dependent genomic adaptations, or shaped species divergence and hybridization histories. Because speciation depends on rates of processes such as genetic mutation, gene flow, and recombination, characterizing rates of landscape change is important for understanding whether certain landscapes facilitate speciation more than others [28] (Figure 2).

Reconstructing climate change

Climate is an important driver of species diversification and adaptation [29–36]. Changes in climate occur over timescales from hundreds to millions of years and are controlled by variations in solar output, Earth's orbit around the sun, and changes in atmospheric concentrations of greenhouse gases. Regional rainfall variations may be influenced by **orographic** effects produced by interactions of water-bearing air masses and topography, with rain shadows commonly forming on the lee side of large mountain ranges. Biologically relevant examples of climate change include variations in annual or seasonal temperature and precipitation, exposure or flooding of continental shelves due to **glacio-eustatic sea-level changes** [37–39], and changes in ocean circulation that shift the location, timing, and seasonality of monsoonal precipitation [40–42]. The formation and maintenance of **climate dipoles**, or two contrasting but adjacent regional climatic conditions, are particularly important because they exert different selection pressures in proximity that can lead to **differential adaptation** (e.g., through **selective sweeps**) or **temporal isolation** of species or populations [43,44]. For example, giraffe lineages in Africa partly diversified due to temporal differences in rainfall-mediated green-up timing caused by precession cycles [45]. Climate fluctuations, such as Northern Hemisphere glacial–interglacial cycles, shift the distribution of habitats and lead to migration and fragmentation of populations. Such events can cause genetic divergence and **bottlenecks** that leave characteristic genomic signatures. Expansion of populations from reduced patches can lead to demographic expansions or increased **mutation burden** in individuals at the front of the shifting or expanding population [46,47].

Prehistoric climate is reconstructed using paleoclimate indicators or proxies such as oxygen isotopes in microfossils (e.g., foraminifera) and marine corals, CO_2 gas preserved in deep ice cores (e.g., on Greenland and Antarctica), growth patterns in tree rings, and varying composition of lake and ocean sediments [48–51]. Paleoclimate data are most useful in geogenomic research if they meet two requirements. First, the timing of past climate change and genetic divergence should be determined independently to assess their synchronicity. Second, if using global datasets it is important

Thermochronology: the study of the thermal history of a mineral, rock, or geologic formation.
Transdisciplinary: the sharing of intellectual knowledge or frameworks beyond or across individual disciplines.
Whole-genome sequence (WGS): sequencing data that covers the entire length of an individual's genome, as opposed to parts of the genome.

Table 1. Summary of disciplines relevant to Earth–life research, their scales, and types of information gained from each

Scientific discipline		Technique/data	Spatial scale	Temporal scale	Knowledge gained	Applications to Earth–life science
Biological	Landscape genetics/genomics	Genetic data (fast-evolving) (e.g., allele frequencies)	Local	Micro	Migration/gene flow estimates, population structure, effect of barriers and local spatial adaptation	Reveals genetic connectivity with recent human-built and natural landscape features
	Phylogeography	Genetic data (mid-evolving) (e.g., sequence data)	Local to regional	Micro–meso	Identity and distribution of genetic groups on the landscape and their relatedness; historical demography	Reveals patterns of genetic diversity across and between physiographic features
	Geogenomics	Genome-scale sequences (e.g., regions of genomic divergence between populations)	Regional	Meso	Evidence for differential adaptation and neutral divergence, historical demography, population relatedness	Reveals mechanisms of genetic divergence across and between physiographic features
	Biogeography	Species distributions	Regional to global	Meso–macro	What factors limit species distributions and how these are shared across species	Associates species distribution patterns with physiographic features
	Phylogenetics	Genetic data (slow-mid evolving) (e.g., phylogenetic trees)	Regional to global	Meso–macro	Relatedness of lineages; whether diversification was concurrent across lineages	Reveals whether lineage divergence and land scape changes occurred at the same time
	Macroecology	Species occurrences	Regional to global	Micro–macro	What regions/features have the highest number of species, among which taxonomic groups, and what variables co-associate.	Associates levels of species richness with physiographic features
Geological/climatic	Geochronology	Radiometric dating (U/Pb, Ar/Ar, ^{14}C), luminescence dating, surface exposure, and cosmogenic nuclide dating (^{10}Be, ^{26}Al, ^{3}He)	Local to global	Micro–macro	Time of rock formation, sediment deposition, surface formation, or exposure	Provides absolute age dates for a geologic event or landscape change
	Thermochronology	Fission track dating, (U-Th)/He dating	Local to global	Micro–macro	Thermal history of a mineral	Characterizes burial and/or exhumation history of minerals, rocks, mountain ranges, or the thermal history of a magma chamber
	Paleoaltimetry	Pollen assemblages, isotopes (δ^{18}O, clumped, volcanic glass)	Local to global	Micro–macro	Past elevation of rocks or geologic formations during formation or deposition	Quantifies the paleo-elevations of landscapes
	Geo-tectonic mapping	Field-based and remote-sensing observations	Local to Global	Micro–macro	Location of and relation between geologic units, fault motions	Determines how a landscape has changed (e.g., faulting, erosion, burial, uplift, volcanic eruptions)
	Sedimentology/ stratigraphy/ geochemistry	Field-based, lab-based, and remote-sensing observations	Local to global	Micro–macro	Ancient sedimentary environments (e.g., terrestrial vs aquatic, alluvial fan vs meandering river)	Provides paleo-environmental interpretations, correlations, and relative age of deposits.
	Volcanology/ geochemistry	X-ray fluorescence, optical mineralogy, stratigraphic study	Local to global	Micro–macro	Style and explosivity of volcanic eruptions and characteristics of an igneous rock	Reveals the timing, duration, and characteristics of volcanic eruptions
	Paleoclimatology/ paleoceanography	Pollen records, tree rings, packrat middens, isotopes in ice cores and marine fossils	Local to global	Micro–macro	Characteristics about paleoenvironmental conditions	Constrains past climate conditions that define habitats and are relevant to the distribution and connectivity of species

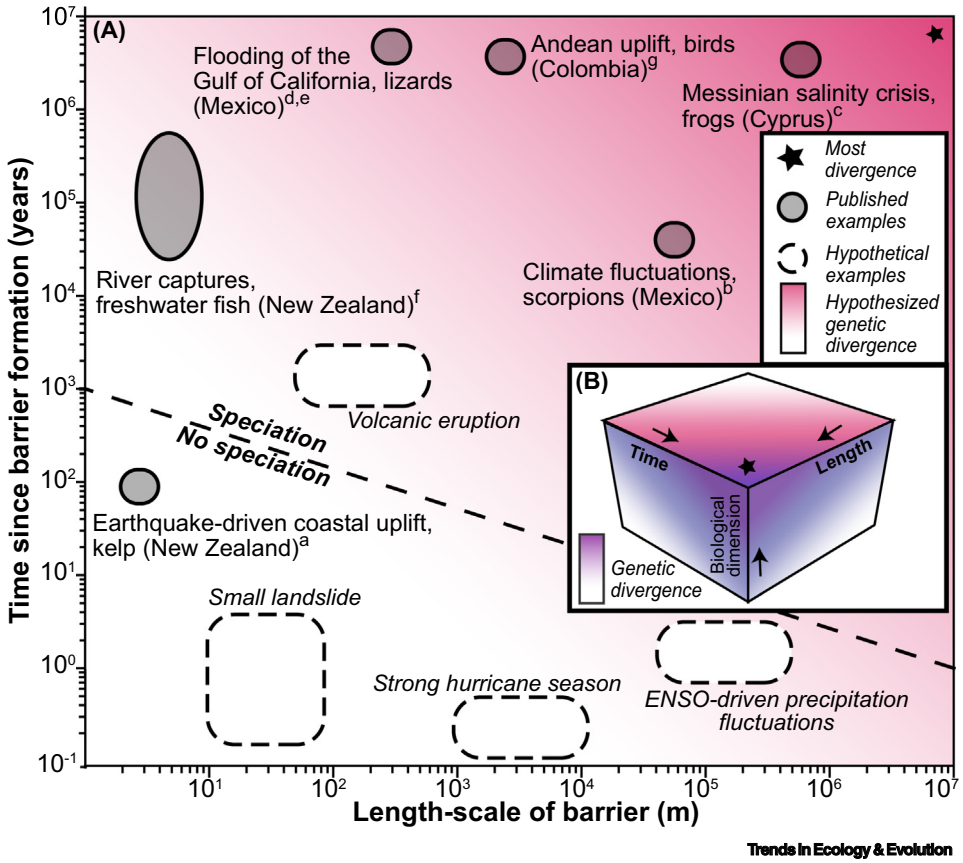

Figure 2. Plot of genetic divergence resulting from different types of geologic and climatic changes that occupy a variety of spatial and temporal scales. **(A)** Hypothesized parameter space relating the time since formation of a barrier to gene flow (y axis), the length-scale of the barrier, (x axis) and the genetic divergence between populations affected by the barrier (magenta color gradient, darker is higher divergence). Gray ovals with black outlines are approximate spatiotemporal extents of published examples of Earth-process-driven genetic divergence [a][96], [b][97], [c][98], [d][13], [e][99], [f][100], [g][101]. White ovals with broken black outlines represent hypothetical Earth processes that can drive genetic divergence. The black dashed line represents a hypothesized threshold below which an Earth process would not drive speciation. **(B)** Conceptualization of how biological dimensions not shown in panel A (e.g., dispersal or generation time) may also influence divergence.

to recognize the potential discrepancies between global and regional changes in rainfall, temperature, and seasonality. These goals can be met by integrating geochronologic data with **general circulation models** that reveal how global climate patterns manifest at regional scales [52–54].

Characterizing evolutionary patterns

Analyzing **whole genome sequence (WGS)** data of populations enables testing detailed biological hypotheses based on geologic and climatic data, as outlined earlier, using the statistical power from the large number of available loci. While loci from reduced representation sequencing methods [e.g., restriction-site associated DNA (RAD), ultra-conserved elements (UCEs)] are suitable for assessing neutral population structure and demography, they only consider a small subset of the genome and therefore are less suitable for testing hypotheses of differential adaptation in which it is critical to know the full range of genes diverging between populations to evaluate cause–effect relationships. For organisms with large genomes, however, neutral loci from reduced representation methods can be paired with transcriptome sequencing to achieve the same goal. Also, by identifying candidate genes within **genomic islands of differentiation** that associate

with ecological functions, these regions can be subsequently studied through more intensive sampling of individuals across the landscape using more cost-effective targeted approaches.

The explicit study of both **neutral elements** and **functional elements** of the genome can be important for differentiating among multiple causal hypotheses. It can, for instance, distinguish the anticipated evolutionary effects of a climate dipole versus those of a river barrier because the former would be expected to produce a signal of differential adaptation near genomic regions under selection. By contrast, the latter might be expected to produce a signal of more neutral population structure observed across the whole genome. In a more complex example, a climate dipole may manifest several effects such as a gradient in precipitation (Figure 3), mean annual temperature, and UV exposure. By sequencing and comparing the genomes and full gene complement of individuals from populations across two adjacent settings, we can use functional enrichment tests (e.g., via **gene ontology**, [55]) of diverged genes to determine whether there is evidence for differential adaptation in genes related to water balance [e.g., aquaporins (AQPs), solute-carrier genes (SLCs)] or in DNA repair and other UV-related genes [e.g., ultraviolet resistance genes (UVRs), DNA damage and repair factors (XPC)] [56], or heat-tolerance genes [e.g., heat shock factors (HSFs), heat shock proteins (HSPs)], or response to different dietary stressors such as toxic plants [e.g., reporter transport proteins (RTPs), taste receptors, type 2 (TAS2Rs)] [57]. One can imagine such an analysis might show that species diverged in response to different aspects of the same climate phenomenon, and paleoclimate proxies and models could be used to estimate when differential selection pressures began, providing putative age control on adaptive phenotypes.

It is important to test potentially pseudocongruent geogenomic hypotheses separately for two reasons. First, while parsimony is useful, it is often plausible that more than one aspect of landscape change has shaped the evolution of an organism. Testing hypotheses independently allows for multiple scenarios to be supported. Second, testing hypotheses separately allows for the possibility that species may evolve in response to different aspects of landscape change. Different responses could be due to life-history traits or niche breadth; such biological differences can be used to characterize higher-order patterns that reveal what groups of organisms are impacted by which aspects of geologic or climatic changes and the limits of these relationships (Figure 3). This approach aims to incorporate into studies the biological, geologic, and climatic complexity inherent in the natural world [3,58].

Importantly, understanding cause–effect relationships within geogenomics also requires surveying the low end of the divergence spectrum — the species that do not show differentiation in response to landscape changes when other species do. These are important negative controls that can be used to characterize and quantify the limits to which geologic and climatic changes affect species. While finding no population structure is often a null or less interesting result, it is critical to understanding how biotic and abiotic processes interact to affect regional evolutionary patterns. Sampling the low end of the divergence spectrum is therefore useful for estimating thresholds and boundary conditions that govern diversification of species within the Earth–life system (Figure 2).

Integrating Earth–life datasets

Geogenomic investigations require a wide variety of data types to reconstruct Earth and evolutionary histories that must be quantitatively integrated. Several statistical and modeling tools can test such hypotheses. Most relevant are modeling approaches that allow explicit, comparative testing of evolutionary scenarios, such as approximate Bayesian computation (e.g., DIYABC2) [59]. These approaches use geologic and/or paleoclimate constraints (e.g., age and duration of a physical barrier, distribution of glacial habitats) with the expected evolutionary result

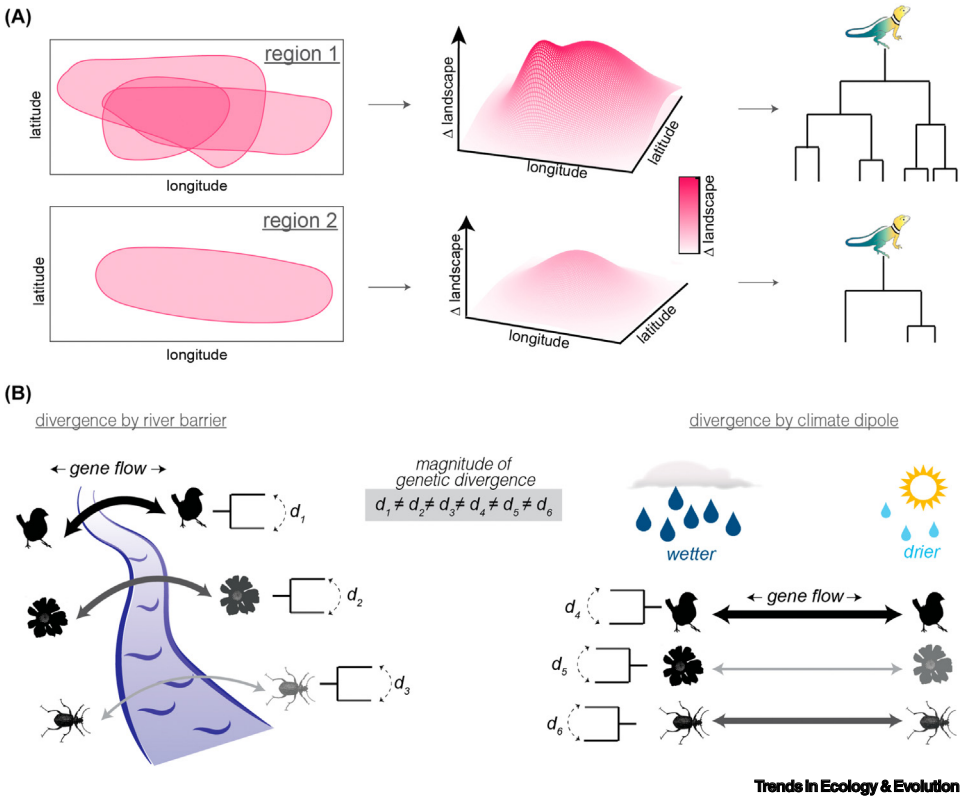

Figure 3. A schematic of outstanding geogenomic questions about the roles of geological and biological complexity within Earth–life systems at the mesoscale. **(A)** Regions with multiple long-lived geological and climatic processes modulating the landscape (region 1) may lead to greater landscape change (middle) and therefore generate higher rates of genetic diversification and speciation (right) than areas with fewer processes (region 2). **(B)** Variations in organismal traits (e.g., dispersal) among species will result in different magnitudes of population divergence (*d*) in response to the formation of a river (left) compared with a climate dipole (right). Color and width of arrows are proportional to gene flow hypothesized for different species. The question of how to quantify and directly compare different types of landscape changes or populational divergence across distantly related species are open questions that require new analytical tools (see Outstanding questions).

(e.g., reduction of gene flow) to simulate genetic data using biological properties of the empirical data [39,60–63]. Statistical comparison of simulations with empirical datasets then reveal which scenarios are best supported. This approach requires careful thought to properly parameterize the anticipated evolutionary effects of geologic and climatic scenarios and simulations can be computationally demanding.

Redundancy analysis [64] can test for correlation between genetic and environmental covariate data (e.g., temperature, precipitation) but can be confounded by collinearity of variables [65–67]. Tree-based software can test for diversification by environment compared to null models [68] and test among biogeographic scenarios [69]. Tools such as Ecoevolity [70] can assess whether populations diverged at the same time, while other tools reveal whether populations diverged without gene flow, with initial gene flow, or with ongoing gene flow (e.g., δaδi) [71]. It is important, though, to consider situations where evolutionary signals might lag behind landscape change, or why species may respond at different rates, for example due to differences in generation time.

Hypotheses of differential adaptation can be tested using sliding window analysis of divergence statistics [72,73] or outlier scans (e.g., Bayescan) [74] to find potentially adaptive loci, and the

biological function of these loci can be statistically summarized using gene enrichment analysis (e.g., gProfiler) [75]; differential adaptation can even be tested in a phylogenetic context across cohorts of species [76]. Finally, forward-simulations of genetic data (e.g., SLiM; [77]; Geonomics, [78]) can be used to simulate the evolutionary effects of landscape changes, but are often more useful for rejecting hypotheses than directly testing them. For example, simulations can help determine if a barrier to gene flow was too ephemeral or occurred too long ago to produce an observed genetic pattern, allowing the rejection of hypotheses without collecting new data. Each approach discussed in this section meets a specific need. Perhaps the greatest current challenge to geogenomic research is insufficient quantitative tools and theory to analyze primary data structures from the Earth and life sciences together. This is the biggest area for growth and will undoubtedly change with time (see following section).

Facilitating convergence science

Future scientific breakthroughs are possible with a geogenomic approach (see Outstanding questions). Such advances will require new ways to facilitate the transfer of knowledge between the Earth and life sciences. This knowledge integration requires the adoption of ways of seeing the world different from our own, or a **transdisciplinary** approach. Within this framework, researchers learn to approach the questions of their field by seeing them from the angles of their collaborators. Geologists learn to see the biological result of geological features and biologists see beyond their study species to the dynamic, physical landscape beneath. There is no substitute for shared discovery, particularly through joint fieldwork. Although it is a slower approach to discovery, creating time and space for this form of collaboration can lead to richer outcomes [105]. One way to facilitate such collaboration is the development of regional working groups that bring together Earth and life scientists studying the same region (Box 1). Such groups provide an interdisciplinary environment to train students in convergence science, and the resulting data products can be adapted for use in science, technology, engineering, and math (STEM) curricula. Physiographic features (e.g., Andes Mountains; [7,8,79–82]), biological patterns (e.g., genetic divergence; Baja GeoGenomics consortium; Box 1), and geographic entities (e.g., New Zealand; [83–87]) are example themes around which groups might assemble. Another option is to facilitate the publication of crosscutting research articles. Annual reviews and regional Earth–life syntheses [7,13,88–90] provide overviews that can be written for a broad audience. Cross-discipline knowledge can also be shared with short editorials that translate the significance of disciplinary findings from one field to another. Today, these are commonly published within a journal, but they could be designed to connect Earth and life science journals within the same publisher. For example, it was long thought that the Isthmus of Panama emerged ~3 Ma, but recent work proposed an alternative model of ~11 Ma [91]. This date is important for biological studies that used the 3-Ma age to constrain divergence times of Pacific and Atlantic sister species. A short summary article could communicate these primary geologic findings into take-home messages relevant for biologists.

Perhaps the biggest advance in the field of geogenomics will come from the development of better tools to model and quantify Earth–life relationships, particularly over timescales that range from sub-thousand-year (e.g., landscape genetics) to million-year (phylogenetics) and in between (geogenomics) [92]. For example, causal structures, which represent cause–effect hypotheses as network graphs, can map relevant relationships at a conceptual level to encourage cross-discipline collaboration [58]. The cause-effect pathways in these models can then be quantified using Bayesian networks or structural equation modeling and applied in a standardized way across species or geographic settings [93–95]. New ways to quantify divergence of populations that can be directly compared across distantly related species would help measure the variability of species responses to the same landscape change, for example resulting from the interaction of abiotic and biotic processes (Figure 3). Finally, quantifying landscape change in biologically meaningful ways

that can be compared across physical settings would be transformational (Figure 3A). We are confident that incentivizing transdisciplinary fields like geogenomics will allow the scientific community to meet these challenges.

Concluding remarks

Many breakthroughs await the integration of Earth and evolutionary sciences at the mesoscale where incipient speciation occurs. The reciprocal use of cutting-edge tools to reconstruct past changes in geology and climate, combined with genome-scale data of populations, allows a geogenomic approach to answer many emergent questions (see Outstanding questions) about the types of physical settings that generate biodiversity, and the limits and conditions under which they operate across taxonomic groups. Darwin called the formation of new species on Earth the 'mystery of mysteries'. Through theoretical, empirical, and collaborative advances, geogenomics can shed new light on the external mechanics, controls, and boundary conditions of this phenomenon.

Acknowledgments

We thank Melanie Culver, Michael Darin, Shanon Fehlberg, Brian Hausback, Kenro Kusumi, and the entire Baja GeoGenomics (BGG) consortium for valuable feedback and ongoing contributions. We also thank Karl Wegmann and two anonymous reviewers for valuable feedback that improved this manuscript. This work was funded by National Science Foundation (NSF) grants EAR-1925535 to G.A.D., EAR-1925560 to R.J.D., and EAR-1925771 to B.T.W. and A.M.-V.. S.E.K.B. was supported by funding from the U.S. Geological Survey National Cooperative Geologic Mapping Program. M.F.S. was supported by the Gaylord Donnelley Postdoctoral Environmental Fellowship from the Yale Institute for Biospheric Studies. A.L.-N. was funded by the Mexican Council for Science and Technology (CONACYT project Ciencia Básica 288690). Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the US Government.

Declaration of interests

No interests are declared.

References

1. Martínez, A. and Mammola, S. (2021) Specialized terminology reduces the number of citations of scientific papers. *Proc. R. Soc. B.* 288, 20202581
2. von Humboldt, A. (1851) *Cosmos: a Sketch of a Physical Description of the Universe, Volume 3*, Harper & Brothers
3. Krishtalka, L. and Humphrey, P. (2000) Can natural history museums capture the future? *BioScience* 50, 611–617
4. McGill, B.J. (2019) The what, how and why of doing macroecology. *Glob. Ecol. Biogeogr.* 28, 6–17
5. Marquet, P.A. (2009) Macroecological perspectives on communities and ecosystem. In *The Princeton Guide to Ecology*, pp. 386–394, Princeton University Press
6. Gaston, K.J. (2000) Global patterns in biodiversity. *Nature* 405, 220–227
7. Hoorn, C. *et al.* (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* 330, 927–931
8. Rahbek, C. *et al.* (2019) Building mountain biodiversity: geological and evolutionary processes. *Science* 365, 1114–1119
9. Rahbek, C. *et al.* (2019) Humboldt's enigma: What causes global patterns of mountain biodiversity? *Science* 365, 1108–1113
10. Lawton, J. (1999) Are there general laws in ecology? *Oikos* 84, 177–192
11. Baker, P.A. *et al.* (2014) The emerging field of geogenomics: constraining geological problems with genetic data. *Earth-Sci. Rev.* 135, 1–10
12. Dickinson, W.R. and Lawton, T.F. (2001) Carboniferous to Cretaceous assembly and fragmentation of Mexico. *Geol. Soc. Am. Bull.* 113, 1142–1160
13. Dolby, G.A. *et al.* (2015) Assessing the geological and climatic forcing of biodiversity and evolution surrounding the Gulf of California. *J. Southwest* 57, 391–455
14. Wan, T. *et al.* (2021) Differences in quaternary co-divergence reveals community-wide diversification in the mountains of southwest China varied among species. *Proc. R. Soc. B* 288, 20202567
15. Antonelli, A. (2017) Biogeography: drivers of bioregionalization. *Nat. Ecol. Evol.* 1, 1–2
16. Cole, T.L. *et al.* (2019) Mitogenomes uncover extinct penguin taxa and reveal island formation as a key driver of speciation. *Mol. Biol.* 36, 784–797
17. Leprieur, F. *et al.* (2016) Plate tectonics drive tropical reef biodiversity dynamics. *Nat. Commun.* 7, 11461
18. Pellissier, L. *et al.* (2018) Are global hotspots of endemic richness shaped by plate tectonics? *Biol. J. Linn. Soc.* 123, 247–261
19. Rabosky, D.L. *et al.* (2018) An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* 559, 1–20
20. Schluter, D. and Pennell, M.W. (2017) Speciation gradients and the distribution of biodiversity. *Nature* 546, 48–55
21. Smith, B.T. *et al.* (2014) The drivers of tropical speciation. *Nature* 406–409
22. Stevens, G.C. (1989) The latitudinal gradient in geographical range: how so many species coexist in the tropics. *Am. Nat.* 133, 240–256
23. Bahadori, A. *et al.* (2018) Reconstruction modeling of crustal thickness and paleotopography of western North America since 36 Ma. *Geosphere* 14, 1207–1231
24. Bennett, S.E.K. *et al.* (2016) Animated tectonic reconstruction of the Lower Colorado River region: implications for Late Miocene to present deformation. In *Going LOCO Investigations along the Lower Colorado River* (Reynolds, R.E., ed.), pp. 73–86, California State University Desert Studies Center
25. Sundell, K.E. *et al.* (2019) Implications of variable late Cenozoic surface uplift across the Peruvian central Andes. *Sci. Rep.* 9, 1–12

Outstanding questions

Are some Earth processes more impactful for shaping the evolution and speciation of organisms than others? If so, what are the characteristics of geologic and climatic processes that determine their impact? It could be the rate of landscape change: perhaps some processes operate too slowly (e.g., vertical crustal motions) and others too quickly (e.g., weather, landslides) to shape evolution. Or does the magnitude or duration of landscape changes dominate (see Figure 2 in the main text)? Answers require comparing the amount and type of landscape change in one area with another, such as the change produced by the development of a climate dipole to that of river formation (see Figure 3 in the main text) or growth of a mountain range. On what spatiotemporal scales and with what units do we measure and compare features?

Perhaps some geologic or climatic processes are more important due to their intrinsic complexity. Uplift of a mountain range may change the landscape more than incision of a river drainage. Both create topography that can isolate populations, but mountains have more causal pathways to influence biological evolution, such as gradients in temperature and precipitation. Considering complex landscape change, is it possible to tease apart past effects of multiple nested deterministic patterns in genomic data? How important are biotic versus abiotic forces and when are genomic signatures maintained, delayed, or overridden? Finally, what are the roles of chance and non-equilibrium conditions in shaping evolutionary patterns we observe today?

Organismally, which traits explain why some species diversify as others do not when faced with the same change to their landscape or climate? Niche breadth or dispersal ability may be important factors. Do species with larger ranges, and thus greater probability of being exposed to landscape or climate changes, show higher speciation rates? Answering these questions requires the ability to compare aspects of population divergence across distantly related organisms.

26. Kar, N. *et al.* (2016) Rapid regional surface uplift of the northern Altiplano plateau revealed by multiproxy paleoclimate reconstruction. *Earth Planet. Sci. Lett.* 447, 33–47
27. Cassel, E.J. *et al.* (2014) Profile of a paleo-orogen: High topography across the present-day basin and range from 40 to 23 Ma. *Geology* 42, 1007–1010
28. Perrigo, A. *et al.* (2020) Why mountains matter for biodiversity. *J. Biogeogr.* 47, 315–325
29. Thom, G. *et al.* (2020) Quaternary climate changes as speciation drivers in the Amazon floodplains. *Sci. Adv.* 6, eaax4718
30. Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philos. Trans. R. Soc. B* 359, 183–195
31. Hewitt, G. (2000) The genetic legacy of the Quaternary ice ages. *Nature* 405, 907–913
32. Jacobs, D.K. *et al.* (2004) Genes, diversity, and geologic process on the Pacific coast. *Annu. Rev. Earth Planet. Sci.* 32, 601–652
33. Saupe, E.E. *et al.* (2019) Spatio-temporal climate change contributes to latitudinal diversity gradients. *Nat. Ecol. Evol.* 3, 1419–1429
34. De Kort, H. *et al.* (2021) Life history, climate and biogeography interactively affect worldwide genetic diversity of plant and animal populations. *Nat. Commun.* 12, 1–11
35. Erwin, D.H. (2009) Climate as a driver of evolutionary change review. *Curr. Biol.* 19, R575–R583
36. Hua, X. and Wiens, J.J. (2013) How does climate influence speciation? *Am. Nat.* 182, 1–12
37. Dolby, G.A. *et al.* (2020) Global physical controls on estuarine habitat distribution during sea level change: consequences for genetic diversification through time. *Glob. Planet. Change* 187, 103128
38. Dolby, G.A. *et al.* (2018) How sea level change mediates genetic divergence in coastal species across regions with varying tectonic and sediment processes. *Mol. Ecol.* 27, 994–1011
39. Stiller, J. *et al.* (2021) Using ultraconserved elements to track the influence of sea-level change on leafy seadragon populations. *Mol. Ecol.* 30, 1364–1380
40. Barron, J.A. *et al.* (2012) Response of the North American monsoon to regional changes in ocean surface temperature. *Paleoceanography* 27, 1–17
41. Lückge, A. *et al.* (2009) Monsoon versus ocean circulation controls on paleoenvironmental conditions off southern Sumatra during the past 300,000 years. *Paleoceanography* 24, 1–14
42. Schott, F.A. *et al.* (2009) Indian Ocean circulation and climate variability. *Rev. Geophys.* 47, 1–46
43. Zuckerberg, B. *et al.* (2020) Climate dipoles as continental drivers of plant and animal populations. *TREE* 35, 440–453
44. Quintero, I. *et al.* (2014) Asynchrony of seasons: genetic differentiation associated with geographic variation in climatic seasonality and reproductive phenology. *Am. Nat.* 184, 352–363
45. Thomassen, H.A. *et al.* (2013) Regional differences in seasonal timing of rainfall discriminate between genetically distinct East African giraffe taxa. *PLoS ONE* 8, e77191
46. Excoffier, L. and Ray, N. (2008) Surfing during population expansions promotes genetic revolutions and structuration. *TREE* 23, 347–351
47. Gilbert, K.J. *et al.* (2018) Mutation load dynamics during environmentally-driven range shifts. *PLoS Genet.* 14, e1007450-18
48. Burke, K.D. *et al.* (2018) Pliocene and Eocene provide best analogs for near-future climates. *Proc. Natl. Acad. Sci. U. S. A.* 115, 13288–13293
49. Lisiecki, L.E. and Raymo, M.E. (2005) A Pliocene-Pleistocene stack of 57 globally distributed benthic δ 18O records. *Paleoceanography* 20, 1–17
50. Marcott, S.A. *et al.* (2013) A reconstruction of regional and global temperature for the past 11,300 years. *Science* 339, 1198–1201
51. Zachos, J. *et al.* (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292, 686–693
52. Judd, E.J. *et al.* (2020) A dynamical framework for interpreting ancient sea surface temperatures. *Geophys. Res. Lett.* 47, e2020GL089044
53. Bhattacharya, T. and Coats, S. (2020) Atlantic-pacific gradients drive last millennium hydroclimate variability in Mesoamerica. *Geophys. Res. Lett.* 47, e2020GL088061
54. Ting, M. *et al.* (2011) Robust features of Atlantic multi-decadal variability and its climate impacts. *Geophys. Res. Lett.* 38, L17705
55. Gene Ontology Consortium (2008) The Gene Ontology project in 2008. *Nucleic Acids Res.* 36, D440–D444
56. Tossi, V.E. *et al.* (2019) Beyond *Arabidopsis*: differential UV-B response mediated by UVR8 in diverse species. *Front. Plant Sci.* 10, 1–17
57. Rocha, J.L. *et al.* (2021) Life in deserts: the genetic basis of mammalian desert adaptation. *TREE* 36, 637–650
58. Dolby, G.A. (2021) Towards a unified framework to study causality in Earth-life systems. *Mol. Ecol.* 30, 5628–5642
59. Cornuet, J.-M. *et al.* (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30, 1187–1189
60. Razgour, O. *et al.* (2015) Unravelling the evolutionary history and future prospects of endemic species restricted to former glacial refugia. *Mol. Ecol.* 24, 5267–5283
61. Zeng, Y.-F. *et al.* (2018) Phylogeographic patterns of the desert poplar in Northwest China shaped by both geology and climatic oscillations. *BMC Evol. Biol.* 18, 75-14
62. Rocha-Méndez, A. *et al.* (2019) The geography of evolutionary divergence in the highly endemic avifauna from the Sierra Madre del Sur, Mexico. *BMC Evol. Biol.* 19, 237–221–21
63. He, Q. *et al.* (2013) Integrative testing of how environments from the past to the present shape genetic structure across landscapes. *Evolution* 67, 3386–3402
64. Capblancq, T. *et al.* (2018) Evaluation of redundancy analysis to identify signatures of local adaptation. *Mol. Ecol. Resour.* 18, 1223–1233
65. Peterman, W.E. and Pope, N.S. (2021) The use and misuse of regression models in landscape genetic analyses. *Mol. Ecol.* 30, 37–47
66. Gaylord, B. and Gaines, S.D. (2000) Temperature or transport? Range limits in marine species mediated solely by flow. *Am. Nat.* 155, 769–789
67. Mcculloch, G.A. *et al.* (2019) Ecological gradients drive insect wing loss and speciation: The role of the alpine treeline. *Mol. Ecol.* 28, 3141–3150
68. Morlon, H. *et al.* (2016) RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol.* 7, 589–597
69. Condamine, F. *et al.* (2013) Diversification patterns and processes of wingless endemic insects in the Mediterranean Basin: historical biogeography of the genus Blaps (Coleoptera: Tenebrionidae). *J. Biogeogr.* 40, 1899–1913
70. Oaks, J.R. (2018) Full Bayesian comparative phylogeography from genomic data. *Syst. Biol.* 68, 371–395
71. Gutenkunst, R.N. *et al.* (2009) Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* 5, e1000695
72. Miller, S.E. *et al.* (2019) A single interacting species leads to widespread parallel evolution of the stickleback genome. *Curr. Biol.* 29, 530–537.e6
73. Shao, Y. *et al.* (2019) Genomic and phenotypic analyses reveal mechanisms underlying homing ability in pigeon. *Mol. Biol. Evol.* 37, 134–148
74. Foll, M. and Gaggiotti, O. (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180, 977–993
75. Reimand, J. *et al.* (2016) g:Profiler-a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res.* 44, 83–89
76. Allio, R. *et al.* (2021) Genome-wide macroevolutionary signatures of key innovations in butterflies colonizing new host plants. *Nat. Commun.* 12, 1–15
77. Haller, B.C. and Messer, P.W. (2019) SLiM 3: forward genetic simulations beyond the Wright–Fisher model. *Mol. Biol. Evol.* 36, 632–637

78. Terasaki Hart, D.E. *et al.* (2021) Genomics: forward-time, spatially explicit, and arbitrarily complex landscape genomic simulations. *Mol. Biol. Evol.* 38, 4634–4646
79. Flantua, S.G. and Hooghiemstra (2018) Historical connectivity and mountain biodiversity. In *Mountains, Climate and Biodiversity* (Hoorn, C. and Perrigo, A. and Antonelli, A., eds), pp. 171–185, Wiley
80. Antonelli, A. *et al.* (2018) Geological and climatic influences on mountain biodiversity. *Nat. Geosci.* 11, 718–726
81. Sanín, M.J. *et al.* (2016) The Neogene rise of the tropical Andes facilitated diversification of wax palms (Ceroxylon: Arecaceae) through geographical colonization and climatic niche separation. *Biol. J. Linn. Soc.* 182, 303–317
82. Sanín, M.J. *et al.* (2017) Up and down the blind alley: population divergence with scant gene flow in an endangered tropical lineage of Andean palms (*Ceroxylon quindiuense* Clade: Ceroxyloideae). *J. Hered.* 108, 288–298
83. Waters, J.M. *et al.* (2015) Within-river genetic connectivity patterns reflect contrasting geomorphology. *J. Biogeogr.* 42, 2452–2460
84. Craw, D. *et al.* (2017) Biological memory of the first Pleistocene glaciation in New Zealand. *Geology* 45, 595–598
85. Craw, D. *et al.* (2016) Evolution of the Taieri River catchment, East Otago, New Zealand. *N. Z. J. Geol. Geophys.* 59, 257–273
86. Wallis, G.P. *et al.* (2016) Transverse alpine speciation driven by glaciation. *TREE* 31, 916–926
87. Wallis, G.P. and Jorge, F. (2018) Going under down under? Lineage ages argue for extensive survival of the Oligocene marine transgression on Zealandia. *Mol. Ecol.* 27, 4368–4396
88. Manish, K. and Pandit, M.K. (2018) Geophysical upheavals and evolutionary diversification of plant species in the Himalaya. *PeerJ* 6, e5919
89. Mosbrugger, V. *et al.* (2018) Cenozoic evolution of geobioiversity in the Tibeto-Himalayan region. In *Mountains, Climate and Biodiversity* (Hoorn, Carina, ed.), pp. 429–448, John Wiley & Sons
90. Wilson, J.S. and Pitts, J.P. (2010) Illuminating the lack of consensus among descriptions of earth history data in the North American deserts: a resource for biologists. *Prog. Phys. Geogr.* 34, 419–441
91. Montes, C. *et al.* (2015) Middle Miocene closure of the Central American Seaway. *Science* 348, 226–229
92. Antonelli, A. *et al.* (2018) Conceptual and empirical advances in Neotropical biodiversity research. *PeerJ* 6, e5644–e5653
93. Grace, J.B. *et al.* (2012) Guidelines for a graph-theoretic implementation of structural equation modeling. *Ecosphere* 3, art73-45
94. Grace, J.B. (2006) *Structural Equation Modeling and Natural Systems*, Cambridge University Press
95. Igea, J. and Tanentzap, A.J. (2021) Global topographic uplift has elevated speciation in mammals and birds over the last 3 million years. *Nat. Ecol. Evol.* 5, 1530–1535
96. Parvizi, E. *et al.* (2020) The genomic footprint of coastal earthquake uplift. *Proc. R. Soc. B* 287, 20200712
97. Graham, M.R. *et al.* (2014) Late Pleistocene to Holocene distributional stasis in scorpions along the Baja California. *Biol. J. Linn. Soc.* 111, 450–461
98. Plötner, J. *et al.* (2010) Genetic divergence and evolution of reproductive isolation in eastern Mediterranean water frogs. In *Evolution in Action*, pp. 373–403, Springer
99. McGuire, J.A. *et al.* (2007) Mitochondrial introgression and incomplete lineage sorting through space and time: Phylogenetics of crotaphytid lizards. *Evolution* 61, 2879–2897
100. Burridge, C.P. *et al.* (2007) An empirical test of freshwater vicariance via river capture. *Mol. Ecol.* 16, 1883–1895
101. Weir, J.T. and Price, M. (2011) Andean uplift promotes lowland speciation through vicariance and dispersal in *Dendrocincla* woodcreepers. *Mol. Ecol.* 20, 4550–4563
102. Riddle, B.R. *et al.* (2000) Cryptic vicariance in the historical assembly of a Baja California peninsular desert biota. *Proc. Natl. Acad. Sci. U. S. A.* 97, 14438–14443
103. Mark, C. *et al.* (2014) Rift flank uplift at the Gulf of California: no requirement for asthenospheric upwelling. *Geology* 42, 259–262
104. Bhattacharya, T. *et al.* (2018) Ice-sheet modulation of deglacial North American monsoon intensification. *Nat. Geosci.* 11, 848–852
105. Clark, S. *et al.* (2020) 6&6: A transdisciplinary approach to art–science collaboration. *BioScience* 70, 821–829. https://doi.org/10.1093/biosci/biaa076