Linking evolutionary and ecological theory illuminates non-stationary biodiversity

A. J. Rominger[1, 2], I. Overcast[3], H. Krehenwinkel[1], R. G. Gillespie[1], J. Harte[1, 4], and M. J. Hickerson[3]

[1]Department of Environmental Science, Policy and Management, University of California, Berkeley
[2]Santa Fe Institute
[3]Biology Department, City College of New York
[4]Energy and Resource Group, University of California, Berkeley

1 Equilibrium in ecology and evolution

We propose that combining ecological theory and inference of evolutionary and demographic change from genetic data will allow us to understand and predict the consequences of non-equilibrial processes in governing the current and future states of biodiversity. The idea of an ecological and evolutionary equilibrium has pervaded studies of biodiversity both on geologic and ecological time scales, and from global to local scales [1, 2, 3, 4, 5, 6, 7, 8, 9]. Biodiversity theories based on assumptions of equilibrium, both mechanistic [5, 7, 9] and statistical [6, 10, 11] have found success in predicting ahistorical patterns of diversity such as the species abundance distribution [12, 5, 6] and the species area relationship [5, 6]. However, investigation of the underlying dynamics producing these patterns has revealed that the equilibrium assumed by the theories is not realistic [13].

The consequences of non-equilibrium dynamics for biodiversity, from diversification to macroecology to conservation, are not well understood. The need to understand non-equilibrial biodiversity processes comes at a critical time when anthropogenic pressures are forcing biodiversity systems into states of rapid transition [14]. The extent to which ecosystems are governed by non-equilibrial processes has profound implications for conservation, which are only just beginning to be explored. For example whether conservation should focus on conventional preservationist paradigms or adaptive management [15]. Whether biodiversity rapidly and consistently tends toward a steady state also determines how species and communities will respond to global environmental change [14].

Tests of equilibrial ecological theory alone will not allow us to identify systems out of equilibrium, nor permit us to pinpoint the mechanistic causes of any observed patterns indicating non-equilibrial processes. The two shortfalls of ahistorical theory are: 1) even when the theory fits observed static patterns, the implicit dynamical assumptions were wrong so we'd make the wrong conclusion about the equilibrium of the system; 2) when the ahistorical theories don't fit the data, we can't know why unless we have a perspective on the temporal dynamics underlying the generation of those patterns that deviate.

Nobody has tried to evaluate stationary v. non-stationary (or history at all) from SAD data alone...so population and species histories inferred from genetic data are critical. and even when, e.g. phylogenies are included it's done so from an equilibrium perspective (but check the lit on this to make sure)

While the fossil record can elucidate deep time patterns for select, well-fossilized groups [16], and in limited geographic areas and temporal extents yielding good preservation [17], we require an approach that is applicable across taxa, and scales of space and time. Bridging ecological theory with models from phylogenetics has long given us potential general-use tools to gain insight into the dynamics underlying contemporary biodiversity patterns [18, 19, 20], while links from population genetics have been more recently explored [18, 19, 20, 21, 22, 23, 24, 25, 26, 27].

While inference of community dynamics from phylogenetic data has its own challenges, particularly in reliably reconstructing extinction rates [28] and species trees topologies in the face of ancestral admixture [29, 30], its general applicability to all life and the advent of economical methods for producing massive amounts of genetic data, make it a promising

approach. Likewise, inference of aggregate population histories underlying regional ecological dynamics using assemblage-level genomic polymorphism data faces statistical challenges, such as non-identifiability of demographic histories [31, 32, 33] and estimating levels of variability across histories [34]. Nevertheless, the promise of unifying processes underlying patterns of species diversities and abundances with distributions of historical population size trajectories, colonisation times, speciation times and regional patterns of genetic connectivity begs investigation. The timing is particularly apt because increasingly reliably and inexpensive methods permit collection of genomic data across a broad range of non-model species.

We propose that combining ecological theory and insights on population and diversity trajectories inferred from phylogenetics and population genetics will allow us to understand and predict the consequences of non-equilibrial processes in governing the current and future states of biodiversity. We argue that this will be achieved by the joint assessment of ahistorical equilibrium patterns and their underlying dynamics using a synthesis of ecological theory, inference techniques derived from phylogenetics and population genetics, and advances in both the wet and dry lab that will unlock massive data resources to test these theories at unprecedented scales. Here we develop this research program by lightly reviewing ecological theory and genetic inference methods, attempts to date at their synthesis, predictions for how non-equilibrial processes will be illuminated by their full synthesis, and prospects for future development, including needed advances in bioinformatic approaches.

1.1 Ecological views on equilibrium and non-equilibrium

The development of the equilibrium theory of island biogeography (ETIB; [35]) ushered in the advent of mechanistically elegant, predictive theories of general patterns in biodiversity. The theory of MacArthur and Wilson also set the precedent of focusing on equilibrial predictions for biodiversity, instead of transient states. From this starting point, three classes of ecological theory have emerged, mechanistically niche-based theory, mechanistically neutral theory, and mechanistically agnostic, statistical mechanical theory.

1.1.1 Mechanistically niche-based theory

Niches have long played a dominant role in guiding ecological theory (cite). The inherent high dimensionality of the niche hypothesis has been recently cast into a lower-dimensional problem by focusing on how differences in resource use (Hutchinsonian niche) and differences in fitness (an aspect of Simpson's adaptive landscape, itself a niche concept) determine competitive coexistence (cite). Competitive coexistence is itself a prediction of equilibrium.

1.1.2 Mechanistically neutral theory

Neutral theory (cite) is another answer to the high dimensionality of the niche hypothesis. By presuming that populations do not differ in fitness nor in resource use, neutrality avoids the intractability of over parameterization and arrives as an equilibrial prediction when

homogeneous stochastic processes of birth, death, speciation and immigration have reached stationarity. Thus neutrality in ecology is analogous to neutral drift in population genetics.

1.1.3 Statistical theory

Rather than assume that any one mechanism, be it niche-based or neutral, dominates the assembly of populations into a community, theories based on statistical mechanics assume that all mechanisms could be valid, but their unique influence has been lost to the enormity of the system and thus outcome of assembly is a community in statistical equilibrium. In one class of such theories, it is assumed that whatever mechanisms are at play, they are only relevant in determining the values of ecological state variables, and then if the system is allowed to come to equilibrium its properties will be predicted by maximizing information entropy relative to the constraints of the state variables. One example is the maximum entropy theory of ecology (METE), one model realization of which assumes that the area (A_0) of an ecosystem, the total number of species (S_0) in some taxonomic group, the total number of individuals in those species (N_0), and the total metabolic rate of those individuals (E_0), capture all necessary information to characterize a community because that community has reached a statistical equilibrium in which the imprint of specific mechanistic forces has been lost. While this theory finds widespread success in predicting ahistorical patterns of species abundance, size, and spatial distribution [6, 12, 36, 37] at single snapshots in time, it fails to match observed patterns in disturbed and rapidly evolving communities [38, 6].

NEED A CLEAR STATEMENT AROUND HERE ABOUT WHAT INSIGHTS WE NEED FROM POP GEN AND PHYLO stemming from the realization that abundance data alone can't identify equilibrial from non-equilibrial systems

1.2 Population genetic and phylogenetic views on equilibrium and non-equilibrium

1.2.1 Population-level inference

Inference of community dynamics that accounts for non-equilibrium historical complexities needs to expand empirical dimensions beyond species abundances and diversities to include axes of information that are historically dynamic with respect to generative models that link spatial-temporal processes and regional genetics patterns across and between species. Coupling phylogenetic and ecological models with extant taxa have been useful [39, 40, 41], yet these approaches lack a link between community-level processes and within-species variation, which could reveal and exploit information about aggregate population histories that underlie non-equilibrium ecological models. Population genetics and phylogeography have always had enormous potential for the inference of community expansion and assembly in the general context of post-LGM warming [42], as well as the community assembly of islands. Yet these studies suffer from the high uncertainty surrounding limited numbers of genetic markers [43], overly generic models [44], ignoring spatial processes [45], and confounding effects of spatial patterns of adaptation [46].

Genomic-scale data, however, is now cost efficient and feasible to collect across a wide swath of non-model species and much progress has been made in multi-species historical demographic models [47]. Although spatial methods are also advancing [48, 49, 50, 51], as are ever more powerful historical inferential approaches via genome-scale data [52], testing community-scale hypotheses with multi-taxa data would be profoundly improved and enriched if population genetic models were grounded in macroecological and biogeographic theory. Conversely, it has been long recognized that models in community ecology have been overly reliant on species abundance distributions which are by themselves often insufficient for distinguishing competing models of assembly without adding other dimensions of data [53, 54].

In contrast to the theoretical developments in ecology that have often focused on geographical and environmental processes underlying the equilibrium dynamics of aggregate species distributions and regional patterns of diversity, the field of phylogeography has taken a historically-centered and species-specific approach by using geographically contextualized population genetic data for testing hypotheses regarding the causal relationship among geographic phenomena, species distributions, and the mechanisms driving speciation. While the development of ecological theory has generally focuses on predictions of abundances and distributions under stationarity, the dynamic nature of landscape and habitat change suggests that ecological theory could be greatly enriched by building a joint modeling framework with population genetic theory that explicitly accounts for historical changes in populations and does not rely on stationarity for generative model predictions.

1.2.2 Coalescent-based inference

One of the fundamental tools allowing for complex historical inference with population genetic data is coalescent theory [55, 56, 57, 58, 59]. Now broadly applied, coalescent theory can generate the statistical properties of any sample of alleles across the genome by modeling gene genealogies backwards in time under virtually any complex demographic history thereby allowing model-based estimation of historical parameters such as historical population size fluctuations, divergence and/or colonization times, and migration rates [60].

Estimating isolation, divergence and/or speciation times has been a particularly important application of population genetic data, and use of coalescent theory is of notable importance in this endeavor because it statistically captures the stochastic discord between population divergence times and gene divergence times [61, 62]. However, the isolation of ancestral lineages into sibling lineages is often only part of a more complex history, as migration and admixture at parts of the genome between diverged populations is a common feature across the tree of life [63, 29, 33, 64], although the frequency and statistical identifiability of this general observation remains highly contentious [65, 66]. In the context of island biogeography and invasion ecology, coalescent-based estimates of isolation times is of particular importance for understanding the dynamics and timing of island colonization, intra-island speciation, as well as invasion times [67, 68, 69, 70].

The history of population size change is also of fundamental importance for understanding the dynamics of community assembly across a variety of ecological settings, and coalescent

theory has likewise become the standard tool for estimating size change histories with population genetic and phylogeographic data on hand [71, 72]. This application of coalescent modeling has been deployed for large numbers of species from which only small numbers of genetic loci are sampled from populations [73], whereas recent advances allowing genome-level data enable far more detailed reconstructions of population history [74, 75] that allow accommodating histories of isolation prior to population size change [52]. However, like any model-based approach, missing assumptions about the complexity of underlying demography can result in biased inference [76], while even using a population history model that matches in reality can not overcome inherent statistical problems in model identifiability [31]. Pivotal to the understanding of demographic and evolutionary histories, coalescent theory has also allowed modeling complex patterns of historical population structure [77, 78], gene flow [79, 80], and even incorporation of extinct "ghost populations" [81, 82] with or without the use of ancient DNA samples [83, 84]. Taking all of these elements of demographic history together (i.e. structure, divergence, expansion, size change and migration), researcher, simulation-based coalescent approaches such as approximate Bayesian computation [85, 86] have become of notable importance for making statistical inference under complex histories when solving the likelihood function becomes intractable [87].

As important as it is for the inference of complex demographic history, coalescent theory has also become an important modeling tool for understanding how natural selection shapes patterns of genetic polymorphism [88, 89, 90]. Indeed, one of the most commonly used techniques for detecting positive selection relies on a summary statistic that can be easily simulated under the coalescent given alternative models with neutrality or selection [91]. However, similar to challenges in ecological theory in distinguishing between neutral and non-neutral models of community assembly with species abundance distributions, population genetic models of positive and/or purifying selection also have very similar predictions to those derived from neutral histories of population growth and/or expansion [92, 93, 94, 95, 96], as well as other more complex models of selection such as polygenic adaptation and interference selection [97, 98]. Regardless of these difficulties, it is likely that various forms of selection and demographic expansion are not mutually exclusive such that making downstream inferences about overall regional biogeographic histories may be accommodated by using coalescent theory to jointly model selection and neutral demographic history to mitigate biased inference [89, 90, 99, 100].

Ultimately, it is at the community level of inference that coalescent-based population genetic methods could be most useful for investigating ecological models that deviate from stationarity. Indeed, it is the inherent historical approach enabled by coalescent methods that can potentially enrich the ecological theoretical approaches to community assembly and stationarity.

In fact, this potential was brought up in the early days of phylogeography [101, 102], as it was well recognized that population genetic data from multiple codistributed taxa could augment investigation of traditionally ecologically-centered questions about the geographic, geological, and/or climatological phenomena that have generated the observed distribution of biodiversity. This proposed "comparative phylogeographic" approach offers the opportunity of a natural experiment where focal objects (codistributed taxa), have been inde-

pendently submitted to the same "natural" evolutionary treatments (geologic and climate change scenarios) [103]. Researchers have generally taken one of two approaches, either by reconstructing taxon-specific histories independently for comparison [104, 105, 106] or using hierarchical statistical models that accommodate aggregate genetic datasets for testing alternative historical scenarios and/or hypotheses at the community level [107, 108, 43, 109].

Despite over 30 years of comparative phylogeographic studies, there has been almost a wholesale neglect of the growing body of theory from community ecology that seeks to accommodate the relative importance of deterministic (e.g., niche filtering, competition) and stochastic (i.e., neutral) processes governing the assembly of communities. Conversely, ecological models of community assembly tend to view communities as static pools with an ahistorical focus on equilibrium expectations. Indeed, a fertile cross pollination of these two bodies of theory could yield a joint inferential framework to bridge together ecological neutral theory with coalescent-based comparative population genetic modeling to better generate predictions of temporal changes in regional patterns of both richness and abundance as well as community-level patterns of genetic diversity and divergence. This whole new type of inference could potentially decouple expectations of abundance distributions from time dependencies by parameterizing the population genetic component of demographic histories underlying temporal changes in abundances.

NEED TRANSITION

2 Current efforts to integrate evolution into ecological theory

Current efforts to synthesize theoretical perspectives from evolution and ecology have made substantial contributions toward understanding what drives biodiversity patterns. However, a more concerted integration is needed, and indeed was not even feasible until recent and ongoing genetic, bioinformatic and theoretical advances. Approaches to date have been hindered by one or more of several general issues: 1) lack of a solid theoretical foundation, 4) inability to distinguish multiple competing alternative hypotheses, 3) lack of comprehensive genetic data, 4) lack of bioinformatic approaches to resolve species and their abundances. Here we quickly survey the ways these shortcomings have prevented further advances and then move on to the cutting edge of the field.

Community phylogenetics [18] attempted to understand the roles of competition and environmental filtering on community assembly by assuming key ecologically-relevant traits are conserved along phylogenies; without a solid theory of trait-mediated competition and recruitment, nor a solid theory of trait evolution, this program broke down [110]. Largely lost is the opportunity to use phylogenetic information to understand the historical contingencies at play in community assembly [111, 112], a task which phylogenies might be able to perform, while they are often poor proxies for traits [110].

Joint studies of genetic and species diversity [24, 113, 114, 26] are largely correlative, lacking a strong theoretical core that could be used to make testable predictions. These

studies also miss the opportunity to explore more than just diversity metrics, but full models of community assembly, population demography and molecular evolution. These studies are also held back by limited access to genetic data, a hurdle we are actively overcoming (see Boxes 1 and 2).

Phylogeographic studies of past climate change have provided insights of how specific groups have responded in non-equilibrium ways to perturbations [103, 104, 106, 109], but such studies cannot make inference about entire community-level processes, nor were they designed to tie into ecological theory and have not been leveraged by theoreticians to gain more realistic insights on demographic and geographic change.

Ecological neutral theory applied to microbial communities [115] has demonstrated that the same ecological processes that operate at macro-scales may also scale down to communities of microbes. However, such studies have not made use of the immense phylogenetic and functional genomic resources available for microbes. Nor has the problem of inferring abundance from metagenomic and metabarcoding data been fully resolved (see Box 2).

2.1   Emerging approaches

Several approaches have been taken that better ground synthesis of ecological and evolutionary dynamics in rigorous theory. Because the NTB is implicitly an evo-ecological theory [5, 116], despite typically being treated as ahistorical, it is natural to include evolutionary information into inference about the theory's parameters. Etienne cast the solution of the NTB's species abundance distribution as a coalescent problem [117] while Jabot and Chave [118] used approximate Bayesian computation to improve estimates of the NTB's fundamental biodiversity number using phylogenetic information. Efforts have also been made to validate the underlying assumption of ecological equivalence, a key assumption of the NTB, from a phylogenetic perspective [119]. While these efforts improved inference of the parameters involved in making ahistorical predictions of species abundance, they did not aim to improve the underlying realism of the evolutionary dynamic presumed by the NTB. For example, while the NTB accurately predicts phylogenetic tree shape (sensu [118]) it does not accurately reflect tree tempo [120]. The time to equilibration in the NTB is also unrealistically long [13]. While protracted speciation has been proposed to correct some of these tempo problems in the NTB [121], it remains to be tested, by a framework such as the one we propose, whether these theoretical advances can accurately predict joint patterns of population genetics, phylogenies, and communities.

Another approach has tested the ahistorical predictions of equilibrial ecological theory through evolutionary snapshots of community assembly and change. Several applications of the NTB in the fossil record have been used to show changes over geologic time in community assembly mechanisms [122, 123]. In a similar theme, Rominger et al. [38] used the geologic chronosequence of the Hawaiian Islands in combination with METE to investigate how evolutionary changes in community assembly drove non-equilibrial patterns in networks of plants and herbivorous insects. While Rominger et al. used genetic information to understand how evolutionary rates vary between different arthropod clades in response to the

geologic chronosequence, these evolutionary snapshot studies lack a quantitative reconciliation of mechanisms inferred by analyses of ahistorical theory with independently inferred dynamics, either from genetic data or stratigraphic time series.

3   What is needed now

A key limitation of using ahistorical theory to infer dynamic mechanisms is that multiple mechanisms, from simple and equilibrial to complex, can map onto the same ahistorical pattern, such as the species abundance distribution [124, 125, 126, 127, 128]. This means that even when a theory describes the data well, we do not really know the dynamics that led to that good fit—an interpretational pitfall common in many studies that claim mechanistic insight even in novel evolutionary study systems [5, 122, 123]. Studies that do not have a strong theoretical foundation, and instead rely on qualitative predictions such as higher or lower phylogenetic dispersion [18], further exacerbate the problem of many mechanisms mapping onto single phenomenological predictions.

Quantitative theoretical foundations and direct information about dynamics can break this many-to-one mapping of mechanism onto theoretical prediction. This nicely parallels calls to incorporate additional information into community ecology and macroecological studies [129]. We propose here a needed framework for integrating the dynamics inferred from population and phylogenetic approaches with with ahistorical, equilibrial ecological theory. There are two complementary options for incorporating the insights of both ahistorical ecological theory and genetic inference methods:

- Option 1: using dynamics from genetic inference to predict and understand deviations from ahistorical theories. This amounts to separately fitting ahistorical theory to typical macroecological data (often species abundance, but potentially including body size and trophic network links) while also fitting population genetic and/or phylogenetic models to genetic data simultaneously captured for the entire community. Doing so requires serious bioinformatic advances that would allow the joint capture of genetic or genomic data from entire community samples, while also estimating accurate abundances for each species in those same samples. This is discussed in Box 2.

- Option 2: building off existing theories, develop new joint models that simultaneously predict macroecological and population genetic patterns. This amounts to building hierarchical models that take genetic data as input and integrate over all possible community states given explicit models of community assembly and population coalescence. Such a model approach also represents a major bioinformatic challenge, which is also discussed in Box 2.

3.1   What we could gain from this framework

Given the insights that could be gained from either option 1 or 2 above, we could finally understand why ahistorical theories fail when they do—is it because of rapid population

change, or evolution/long-distance dispersal of novel ecological strategies? We could predict whether a system that obeys the ahistorical predictions of equilibrial ecological theory is in fact undergoing major non-equilibrial evolution. We could better understand and forecast how/if systems out of equilibrium are likely to relax back to equilibrial patterns. With such a framework we could even flip the direction of causal inference and understand ecological drivers of diversification dynamics. This last point bears directly on long-standing and open debates about the importance of competitive limits on diversification. Competition and limiting similarity have a long history of study as drivers of diversification. This has culminated in ideas of diversity-dependent diversification[130, 131, 132]. What has not been done is link this back to ecological assembly mechanisms, but the opportunity seems ripe considering the abundance of work on niche differences and fitness differences[7, 133, 134, 135]. There has even been work on this from a phylogenetic viewpoint[136, 137]. Conclusions about phylogenetic patterns (e.g. diversification slowdowns) would be more believable and robust if combined with population genetic inference (e.g. declining populations) and community patterns (e.g. deviation from equilibrium).

NEED TRANSITION

4 Evo-ecological predictions for systems out of equilibrium

The data needed to fully test a non-equilibrial theory of ecology and evolution, synthesizing historical and contemporary biodiversity patterns, are unprecedented in scale and depth. Put simply, we require knowing the species identities of each individual in a sample as well as information on some portion of their genomes such that we can estimate historical demography and diversification. In Box 2 we discuss progress toward generating such data. We highlight two promising routes: 1) estimating abundance from targeted capture high throughput sequencing data (i.e. metabarcoding) to be used in ahistorical ecological theory testing, and then separately fitting models of demography and diversification; and 2) jointly estimating the parameters of coupled models of community assembly and population demographics. Assuming these two approaches are within reach (as we demonstrate in **Bioinformatic advances**), we now discuss hypotheses to be tested in our non-equilibrium framework.

4.1 Cycles of non-equilibrium

Ecosystems are likely to experience consistence disturbances (cite) which can occur on ecological time-scales, such as primary success (cite), or evolutionary time scales, such as evolution of novel innovations that lead to new ecosystem processes (cite). We hypothesize that these consistence disturbances will lead to cycles of non-equilibrium in observed biodiversity patterns.

Figure 1 shows how we hypothesize biodiversity to transition between different phases

of equilibrium and non-equilibrium. The figure presents deviation between data and ahistorical ecological theory on the y-axis, versus deviations from demographic or evolutionary equilibrium on the x-axis. A clockwise cycle through this space would indicate:

- Panel I → Panel II: following rapid ecological disturbance, ecological metrics diverge from equilibrium

- Panel II → III: ecological non-equilibrium spurs evolutionary non-equilibrium leading to both ecological and evolutionary metrics diverging from equilibrium values

- Panel III → IV: ecological relaxation to equilibrium after evolutionary innovations provide the means for populations to re-equilibrate to their environments

- Panel IV → I: finally a potential return to equilibrium of both ecological and evolutionary metrics once evolutionary processes have also relaxed to their equilibrium

Cycles could also be much shorter, with a system only transitioning back and forth between Panel I and Panel II. This scenario corresponds to the system being driven only by rapid ecological disturbance, and this disturbance itself following a stationary dynamic leading to no net evolutionary response.

Cycles through this space could also occur in a counterclockwise direction, being initiated by an evolutionary innovation. Under such a scenario we hypothesize the cycle to proceed:

- Panel I → IV: non-equilibrium evolution (including sweepstakes dispersal) leading to departure from evolutionary equilibrium before departure from ecological equilibrium

- Panel IV → III: non-equilibrial ecological response to non-equilibrium evolutionary innovation

- Panel III → I: ecological and evolutionary relaxation

We hypothesize that the final transition will be directly to a joint equilibrium in ecological and evolutionary metrics (Panel I) because a transition from panel III to panel II is unlikely, given that ecological rate are faster than evolutionary rates.

A complete cycle cannot be observed without a time machine, but by combining ahistorical ecological theory and population/phylogenetic inference methods with community-level genetic data we can identify where on the cycle our focal systems are located. Such an approach assumes that abundance data have been estimated from sequence data, ahistorical ecological theories have been fit to those abundance data, and models of population demography and/or diversification have been separately fit to the underlying sequence data. To better under how our focal systems have transitioned between different equilibrium and non-equilibrium phases, we must more deeply explore the joint inference of community assembly and evolutionary processes. In the following sections we do just that for each transition shown in Figure 1. We bring to bear other aspects of joint eco-evolutionary inference, in particular the 1) relationship between lineage age (inferred from molecular data) and lineage

abundance; 2) the nature of deviation from ecological metrics, specifically the shape of the species abundance distribution; and 3) the nature of deviation from evolutionary metrics, specifically inference of past population change and selection.

4.2   Systems undergoing rapid ecological change

For systems whose metrics conform to demographic predictions of equilibrium, but deviate from equilibrial ecological theory, we predict that rapid ecological change underlies their dynamics. However, more information is needed to confirm that the system is being driven primarily by rapid ecological change. The first line of evidence could come from a lack of correlation between lineage age and lineage abundance—this would indicate that slow eco-evolutionary drift is interrupted by frequent perturbations to populations, making their size independent of age (Fig. 2). Actual abundance should similarly be uncorrelated with inference of effective population size from genetic data. Further support for the ecology-only hypothesis could come from a lack of directional selection detected in community-wide surveys of large genomic regions (see Boxes 1 and 2). Taken as a whole, systems in which ecological metrics deviate from equilibrial theory while demographic and macroevolution-ary metrics conform to equilibrial theory presents an opportunity to understand and test hypotheses relating to disturbance, assembly, and the shape of the species abundance distribution (cite).

4.3   Non-equilibrium ecological communities fostering non-equilibrium evolution

A lack of equilibrium in an ecological assemblage means that the system will experience change on its trajectory toward a future possibility of equilibrium. If ecological relaxation does not occur—by chance, or because no population present is equipped with the adaptations to accommodate the new environmental that drove the system from equilibrium to begin with—then the system is open to evolutionary innovation. Such innovation could take the form of elevated speciation or long-distance immigration of populations with different ecological strategies than those present, relating to the idea that community assembly is a race between processes with potentially different, but stochastic rates [24], thus any process might by chance produce the observed outcome. In this way, speciation and sweepstakes immigration (and its human-added counterpart, invasion) represent similar processes. Speciation and sweepstakes/invasion will yield very different phylogenetic signals, however their population genetic signals in a non-equilibrium community may be very similar (e.g. rapid population expansion). Thus where non-equilibrium communities foster non-equilibrium diversification (either through speciation or invasion) we expect to see a negative relationship between lineage age and abundance (Fig. 2) and signs of rapid population expansion in the newest arrivals to the system.

Local extinctions results from non-equilibrium community dynamics are unlikely to scale up to non-equilibrial macroevolution, so long as the extinctions are not global. For such

ecosystem collapse to occur, continental ecosystems would have to be self-organized to the point of criticality, a hypothesis which has largely been refuted [138].

4.4 Non-equilibrium evolution fostering non-equilibrium ecological dynamics

If evolutionary processes, or their counterpart in the form of sweepstakes immigration/invasion, generate new ecological strategies in a community, this itself constitutes a form of disturbance pushing the system to reorganize, thus our predicted cycle in Figure 1 from Panel I to IV to III. Evolutionary change would have to be extremely rapid to force ecological metrics out of equilibrium, because evolutionary change could be accommodated by ecological communities if it occurs slowly enough, thus we would expect to see phylogenetic signals of adaptive radiation, and corresponding signals of strong selection in genomic-scale sequence data.

4.5 Ecological relaxation

Ecological metrics can return to equilibrium either by ecological means (local immigration from the source pool, local extinction, rapid population changes) or by evolutionary means (novel innovations arising through selection and adaptation or sweepstakes immigration). In either case, communities are likely to return to equilibrium given enough time without disturbance. Because ecological rates are typically faster than evolutionary rates, this ecological relaxation is likely to happen more quickly than evolutionary relaxation, and thus genetic inference may reveal a time-averaged demographic signature of non-equilibrium.

4.6 Evolutionary relaxation

Given sufficient time in ecological equilibrium, the time averaged demographic record revealed by genetic inference will likely also re-equilibrate.

5 Harnessing evo-ecological measures of non-equilibrium for a changing world

Conclude on why this all matters.

6 More future directions

1. Integration with fossils beyond dating phylogenies

13

2. Functional genomics, already underway for microbes, but soon to be realistic for macro-organism communities

Glossary and boxes

Glossary

ahistorial equilibrium

Box 1: Wetlab techniques

Metabarcoding describes the targeted PCR amplification and next generation sequencing of short DNA barcode markers (typically 300-500 bp) from community samples (Ji et al. 2013; Yu et al. 2012). The resulting amplicon sequences are clustered into operational taxonomic units (OTUs), based on sequence similarity and compared to reference databases for taxonomic assignment. Even minute traces of taxa in environmental samples can be detected using metabarcoding (Bohmann et al. 2014). At the same time, amplicon sequencing is cheap, requires a small workload and thus allows rapid inventories of species composition and species interactions in whole ecosystems (Gibson et al. 2014; Leray & Knowlton 2015, Pompanon et al. 2012). However, the phylogenetic resolution offered by short barcode markers is limited. And the preferential amplification of some taxa during PCR can lead to highly skewed abundance estimates (Elbrecht & Leese 2015) or the complete loss of some taxa (Yu et al. 2012) from metabarcoding libraries. Recent work (Saitoh et al. 2016) suggests the application of taxon specific read abundance correction factors to account for amplification bias and derive accurate estimates for taxa from amplicon sequencing data.

Metagenomic approaches, in contrast, avoid marker specific amplification bias by sequencing libraries constructed either from untreated genomic DNA (Dodsworth 2015; Linard et al. 2015; Tang et al. 2014), or after targeted enrichment of genomic regions (Liu et al. 2016). While being more laborious, expensive and computationally demanding than metabarcoding, metagenomics thus offers improved accuracy in detecting species composition and abundance (Zhou et al. 2013). Moreover, the assembly of high coverage metagenomic datasets recovers large contiguous sequence stretches, even from rare members in a community, offering high phylogenetic resolution at the whole community level (Coissac et al. 2016). Due to large genome sizes and high genomic complexity, metazoan metagenomics is currently mostly limited to the assembly of fairly short high copy regions. Particularly mitochondrial and chloroplast genomes as well as nuclear ribosomal clusters are popular targets (Dodsworth 2015; Coissac et al. 2016). In contrast, microbial metagenomic studies now routinely assemble complete genomes and characterize gene content and metabolic pathways even from complex communities (Nielsen et al. 2014). This allows unprecedented insights into functional genetic process underlying community assembly and evolutionary change of communities to environmental stress.

Box 2: Bioinformatic advances

Next generation sequencing technology has ushered in a revolution in evolutionary biology and ecology. The large scale recovery of species richness, food web structure, cryptic species, identification of juveniles and hidden diversity, e.g. internal parasitoids, promise unprecedented new insights into ecosystem function and assembly [see Box 1; ? ? ? ?]. While species richness can be routinely identified by sequencing bulk samples, estimating species abundance remains challenging [?] and severely limits the application of high throughput sequencing methods to many community-level studies. We propose two complementary approaches to estimate species abundance from highthroughput sequenc data, while of course retaining the genetic information. The first approach estimates abundance free from any models of community assembly, the second jointly estimates the parameters of a specific assembly model of interest along with the parameters of a coalescent-based population genetic/demographic model. Both methods assume genetic data are availible from a pooled sample of organisms, for example as collected by many pasive sampling methods for arthropod communities or in sampling microbial communities from the environment.

Model-free abundance estimation. We propose a pipline (Fig. I) where raw reads are generated and assembled into a phylogeny using standard approaches, and potentially added by additionally availible sequence data in a super tree or super matrix approache. The numbers of sequences assigned to each terminal tip are then used in a Bayesian hierrarchical model which seeks to estimate the true number of organisms representing each terminal tip while accounting for sequencing biases originating from primer affinity and copy number differences between taxa. Information on phylogenetic relatedness will be used to inform modeled correlations in biases between taxa [? , e.g. copy number is known to be phylogenetically conserved]. This approach is particularely taylored to metabarcoding data. In a potentially powerful approach, and thanks to the proposed Bayesian framework, information from sequencing experiments that seek to calibrate metabarcoding studies [e.g., ?] can be used to build meaningfully informative priors and improve model accuracy. In a simplistic simulation study (described in the supplement) we show that such an approach can accurately estimate true underlying abundances while accounting for known biases (Fig. II).

Joint inference of community assembly and population genetic models. Coupling individual-based, forward-time models of community colonization and expansion with a backwards-time hierarchical multi-taxa coalescent models will allow us to make inference about the values of the parameters in both models. This framework is flexible enough to incorporate multiple refugia, colonization routes, ongoing migration and parameterizing both neutral and deterministic processes of assembly on time scales of hundreds of thousands of years (Fig. III). So far we have implemented a simulation model in the gimmeSADπ [?] by merging an individual-based forward time community dispersal model [?] with the msPrime coalescent simulator [?]. This has been accomplished by rescaling the time dependent local abundance distributions into time- dependent effective population size distributions while allowing for heterogeneity in migration and colonization rates. This simulation model can be combined with random forest classifiers and hierarchical ABC to enable testing alternative assembly models, including models that have not yet reached their theoretical equilibria.
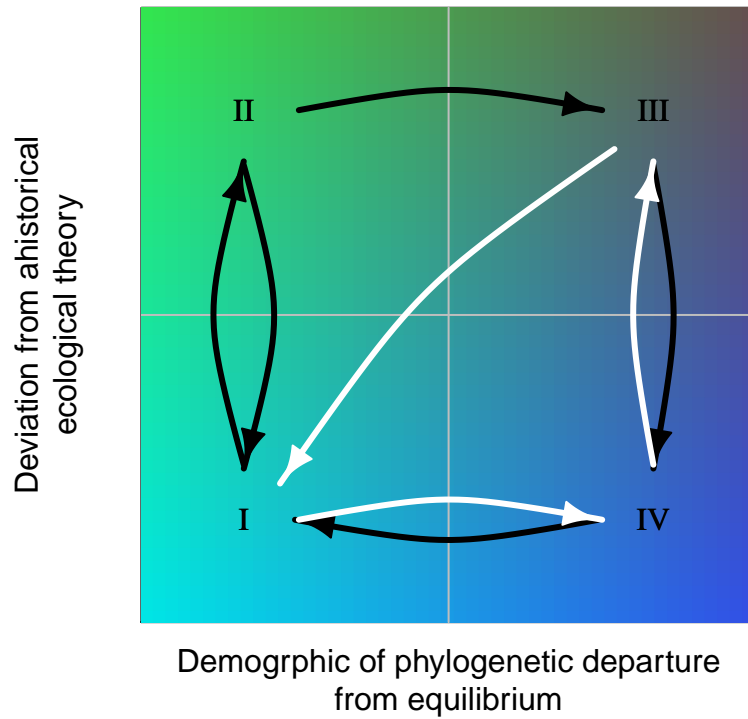
Figures

Figure 1: Hypothesized cycles between different states of equilibrium and non-equilibrium in ecological metrics (y-axis) and evolutionary metrics (x-axis). Panels I–IV are discussed in the next. Colors correspond to deviation from ahistorical ecological theory and evolutionary equilibrium. Black cycle corresponds to non-equilibrium initiated by ecological disturbance (with potential to continue to evolutionary non-equilibrium or relaxation to equilibrium). White cycle is initiated by evolutionary innovation.
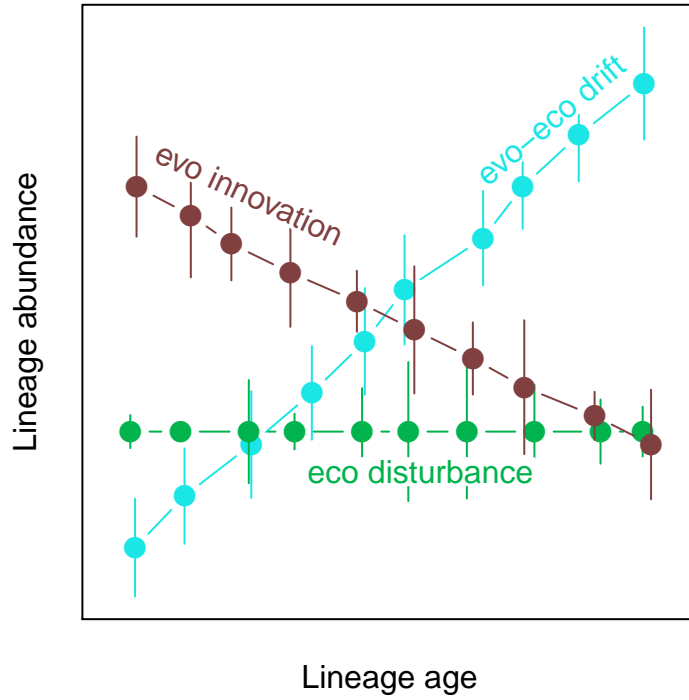
Figure 2: Hypothesized relationships between lineage age and abundance under different evo-ecological scenarios. Colors correspond to panels in Figure 1: teal is evo-ecological equilibrium; green is rapid transition to ecological non-equilibrium following short timescale distrubance; dark brown is non-equilibrium in both ecological and evolutionary metrics.

References

[1] Sepkoski, J.J. (1984) A kinetic model of phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinctions. *Paleobiology* 10, 246–267

[2] Alroy, J. (2010) The shifting balance of diversity among major marine animal groups. *Science* 329, 1191–1194

[3] Rabosky, D.L. and Lovette, I.J. (2008) Density-dependent diversification in north american wood warblers. *Proceedings of the Royal Society B: Biological Sciences* 275, 2363–2371

[4] Rabosky, D.L. (2009) Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecol. Lett.* 12, 735–743

[5] Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*, vol. 32. Princeton University Press

[6] Harte, J. (2011) *The Maximum Entropy Theory of Ecology.* Oxford University Press

[7] Chesson, P. (2000) Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* 31, 343–366

[8] Adler, P.B. *et al.* (2010) Coexistence of perennial plants: an embarrassment of niches. *Ecol. Lett.* , no–no

[9] Tilman, D. (2004) Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. *Proc. Natl. Acad. Sci. U. S. A.* 101, 10854–10861

[10] Pueyo, S. *et al.* (2007) The maximum entropy formalism and the idiosyncratic theory of biodiversity. *Ecol. Lett.* 10, 1017–1028

[11] Shipley, B. *et al.* (2006) From plant traits to plant communities: A statistical mechanistic approach to biodiversity. *Science*

[12] White, E.P. *et al.* (2012) Characterizing species abundance distributions across taxa and ecosystems using a simple maximum entropy model. *Ecology* 93, 1772–1778

[13] Ricklefs, R.E. (2006) The unified neutral theory of biodiversity: do the numbers add up? *Ecology* 87, 1424–1431

[14] Barnosky, A.D. *et al.* (2012) Approaching a state shift in earth/'s biosphere. *Nature* 486, 52–58

[15] Wallington, T.J. *et al.* (2005) Implications of current ecological thinking for biodiversity conservation: A review of the salient issues. *Ecol. Soc.* 10

[16] Alroy, J. *et al.* (2008) Phanerozoic trends in the global diversity of marine invertebrates. *Science* 321, 97–100

[17] Harnik, P.G. (2011) Direct and indirect effects of biological factors on extinction risk in fossil bivalves. *Proceedings of the National Academy of Sciences* 108, 13594–13599

[18] Webb, C.O. *et al.* (2002) Phylogenies and community ecology. *Annu. Rev. Ecol. Syst.* 33, 475–505

[19] Emerson, B.C. (2002) Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Mol. Ecol.* 11, 951–966

[20] Lavergne, S. *et al.* (2010) Biodiversity and climate change: Integrating evolutionary and ecological responses of species and communities. *Annu. Rev. Ecol. Evol. Syst.* 41, 321–350

[21] Li, L. and Ma, Z.S. (2016) Testing the neutral theory of biodiversity with human microbiome datasets. *Sci. Rep.* 6, 31448

[22] McGaughran, A. (2015) Integrating a population genomics focus into biogeographic and macroecological research. *Front. Ecol. Evol.* 3

[23] Laroche, F. *et al.* (2015) A neutral theory for interpreting correlations between species and genetic diversity in communities. *Am. Nat.* 185, 59–69

[24] Vanoverbeke, J. *et al.* (2015) Community assembly is a race between immigration and adaptation: eco-evolutionary interactions across spatial scales. *Ecography*

[25] Vellend, M. and Geber, M.A. (2005) Connections between species diversity and genetic diversity. *Ecol. Lett.* 8, 767–781

[26] Papadopoulou, A. *et al.* (2011) Testing the Species–Genetic diversity correlation in the aegean archipelago: Toward a Haplotype-Based macroecology? *Am. Nat.* 178, 241–255

[27] Dexter, K.G. *et al.* (2012) Historical effects on beta diversity and community assembly in amazonian trees. *Proc. Natl. Acad. Sci. U. S. A.* 109, 7787–7792

[28] Quental, T.B. and Marshall, C.R. (2009) EXTINCTION DURING EVOLUTIONARY RADIATIONS: RECONCILING THE FOSSIL RECORD WITH MOLECULAR PHYLOGENIES. *Evolution* 63, 3158–3167

[29] Mallet, J. *et al.* (2016) How reticulated are species? *Bioessays* 38, 140–149

[30] Xu, B. and Yang, Z. (2016) Challenges in species tree estimation under the multispecies coalescent model. *Genetics* 204, 1353–1368

[31] Terhorst, J. and Song, Y.S. (2015) Fundamental limits on the accuracy of demographic inference based on the sample frequency spectrum. *Proc. Natl. Acad. Sci. U. S. A.* 112, 7677–7682

[32] Robinson, J.D. *et al.* (2014) ABC inference of multi-population divergence with admixture from unphased population genomic data. *Mol. Ecol.* 23, 4458–4471

[33] Sousa, V. and Hey, J. (2013) Understanding the origin of species with genome-scale data: modelling gene flow. *Nat. Rev. Genet.* 14, 404–414

[34] Hickerson, M.J. *et al.* (2014) Recommendations for using msbayes to incorporate uncertainty in selecting an abc model prior: a response to oaks et al. *Evolution* 68, 284–294

[35] MacArthur, R.H. and Wilson, E.O. (1967) *The theory of island biogeography*. Princeton University Press

[36] Xiao, X. *et al.* (2015) A strong test of the maximum entropy theory of ecology. *Am. Nat.* 185, E70–E80

[37] Harte, J. *et al.* (2009) Biodiversity scales from plots to biomes with a universal species-area curve. *Ecol. Lett.* 12, 789–797

[38] Rominger, A.J. *et al.* (2015) Community assembly on isolated islands: macroecology meets evolution. *Glob. Ecol. Biogeogr.*

[39] Emerson, B.C. and Gillespie, R.G. (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends Ecol. Evol.* 23, 619–630

[40] Rosindell, J. *et al.* (2015) Unifying ecology and macroevolution with individual-based theory. *Ecol. Lett.* 18, 472–482

[41] Burbrink, F.T. *et al.* (2015) Predicting community structure in snakes on eastern nearctic islands using ecological neutral theory and phylogenetic methods. *Proc. Biol. Sci.* 282

[42] Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the quaternary. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 359, 183

[43] Chan, Y.L. *et al.* (2014) Detecting concerted demographic response across community assemblages using hierarchical approximate bayesian computation. *Mol. Biol. Evol.* , msu187

[44] Papadopoulou, A. and Knowles, L.L. (2016) Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. *Proc. Natl. Acad. Sci. U. S. A.* 113, 8018–8024

[45] Meirmans, P.G. (2012) The trouble with isolation by distance. *Mol. Ecol.* 21, 2839–2846

[46] Hoban, S. *et al.* (2016) Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *Am. Nat.* 188, 379–397

[47] Xue, A.T. and Hickerson, M.J. (2015) The aggregate site frequency spectrum (aSFS) for comparative population genomic inference. *Mol. Ecol.*

[48] Petkova, D. *et al.* (2016) Visualizing spatial population structure with estimated effective migration surfaces. *Nat. Genet.* 48, 94–100

[49] Joseph, T.A. *et al.* (2016) Demographic inference under a spatially continuous coalescent model. *Heredity* 117, 94–99

[50] Prates, I. *et al.* (2016) Inferring responses to climate dynamics from historical demography in neotropical forest lizards. *Proceedings of the National Academy of Sciences* 113, 7978–7985

[51] Brown, J.L. *et al.* (2016) Predicting the genetic consequences of future climate change: The power of coupling spatial demography, the coalescent, and historical landscape changes. *Am. J. Bot.* 103, 153–163

[52] Terhorst, J. *et al.* (2016) Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nat. Genet.*

[53] McGill, B. (2003) Strong and weak tests of macroecological theory. *Oikos* 102, 679–685

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

[55] Hudson, R.R. (1983) Properties of a neutral model with intragenic recombination. *Theor. Popul. Biol.* 23, 183–201

[56] Tajima, F. (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105, 437–460

[57] Kingman, J.F.C. (1982) The coalescent. *Stochastic Process. Appl.* 13, 235–248

[58] Kingman, J.F.C. (1982) On the genealogy of large populations. *J. Appl. Probab.* 19a, 27–43

[59] Rosenberg, N.A. and Nordborg, M. (2002) Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nat. Rev. Genet.* 3, 380–390

[60] Wakeley, J. (2008) *Coalescent Theory.* Roberts and Company Publishers

[61] Charlesworth, D. (2010) Don't forget the ancestral polymorphisms. *Heredity* 105, 509–510

[62] Edwards, S.V. and Beerli, P. (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* 54, 1839–1854

[63] Shapiro, B.J. *et al.* (2016) What is speciation? *PLoS Genet.* 12, e1005860

[64] Nosil, P. (2008) Speciation with gene flow could be common. *Mol. Ecol.* 17, 2103–2106

[65] Cruickshank, T.E. and Hahn, M.W. (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol.* 23, 3133–3157

[66] Yang, M. *et al.* (2017) Can genomic data alone tell us whether speciation happened with gene flow? *Mol. Ecol.*

[67] Estoup, A. and Clegg, S.M. (2003) Bayesian inferences on the recent island colonization history by the bird zosterops lateralis lateralis. *Mol. Ecol.* 12, 657–674

[68] Estoup, A. *et al.* (2004) Genetic analysis of complex demographic scenarios: spatially expanding populations of the cane toad, bufo marinus. *Evolution* 58, 2021–2036

[69] Hickerson, M.J. and Meyer, C. (2008) Testing comparative phylogeographic models of marine vicariance and dispersal using a hierarchical bayesian approach. *BMC Evol. Biol.* 8, 322

[70] Gray, M.M. *et al.* (2014) Demographic history of a recent invasion of house mice on the isolated island of gough. *Mol. Ecol.* 23, 1923–1939

[71] Kuhner, M.K. *et al.* (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics* 149, 429–434

[72] Slatkin, M. and Hudson, R.R. (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129, 555–562

[73] Drummond, A.J. *et al.* (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* 22, 1185–1192

[74] Schiffels, S. and Durbin, R. (2014) Inferring human population size and separation history from multiple genome sequences. *Nat. Genet.* 46, 919–925

[75] Boitard, S. *et al.* (2016) Inferring population size history from large samples of Genome-Wide molecular data - an approximate bayesian computation approach. *PLoS Genet.* 12, e1005877

[76] Mazet, O. *et al.* (2015) On the importance of being structured: instantaneous coalescence rates and human evolution—lessons for ancestral population size inference? *Heredity* 116, 362–371

[77] Prado-Martinez, J. *et al.* (2013) Great ape genetic diversity and population history. *Nature* 499, 471–475

[78] Bahlo, M. and Griffiths, R.C. (2000) Inference from gene trees in a subdivided population. *Theor. Popul. Biol.* 57, 79–95

[79] Beerli, P. and Felsenstein, J. (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc. Natl. Acad. Sci. U. S. A.* 98, 4563–4568

[80] Hey, J. and Nielsen, R. (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of drosophila pseudoobscura and d. persimilis. *Genetics* 167, 747–760

[81] Slatkin, M. (2005) Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. *Mol. Ecol.* 14, 67–73

[82] Alter, S.E. *et al.* (2007) DNA evidence for historic population size and past ecosystem impacts of gray whales. *Proc. Natl. Acad. Sci. U. S. A.* 104, 15162–15167

[83] Kuhlwilm, M. *et al.* (2016) Ancient gene flow from early modern humans into eastern neanderthals. *Nature* 530, 429–433

[84] Veeramah, K.R. and Hammer, M.F. (2014) The impact of whole-genome sequencing on the reconstruction of human population history. *Nat. Rev. Genet.* 15, 149–162

[85] Beaumont, M. (2010) Approximate bayesian computation in evolution and ecology. *Annu. Rev. Ecol. Evol. Syst.* 41, 379–406

[86] Pritchard, J.K. *et al.* (1999) Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Mol. Biol. Evol.* 16, 1791–1798

[87] Sunnåker, M. *et al.* (2013) Approximate bayesian computation. *PLoS Comput. Biol.* 9, e1002803. doi:10.1371/journal.pcbi.1002803

[88] Kim, Y. and Stephan, W. (2002) Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics* 160, 765–777

[89] Kern, A.D. and Schrider, D.R. (2016) Discoal: flexible coalescent simulations with selection. *Bioinformatics* 32, 3839–3841

[90] Ewing, G.B. and Jensen, J.D. (2016) The consequences of not accounting for background selection in demographic inference. *Mol. Ecol.*

[91] Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595

[92] Freedman, A.H. *et al.* (2016) Demographically-based evaluation of genomic regions under selection in domestic dogs. *PLoS Genet.* 12, e1005851

[93] Barton, N.H. (1998) The effect of hitch-hiking on neutral genealogies. *Genet. Res.* 72, 123–133

[94] Barton, N.H. (2000) Genetic hitchhiking. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 355, 1553–1562

[95] Jensen, J.D. *et al.* (2005) Distinguishing between selective sweeps and demography using DNA polymorphism data. *Genetics* 170, 1401–1410

[96] Schrider, D.R. *et al.* (2016) Effects of linked selective sweeps on demographic inference and model selection. *Genetics* 204, 1207–1223

[97] Stephan, W. (2016) Signatures of positive selection: from selective sweeps at individual loci to subtle allele frequency changes in polygenic adaptation. *Mol. Ecol.* 25, 79–88

[98] Good, B.H. *et al.* (2014) Genetic diversity in the interference selection limit. *PLoS Genet.* 10, e1004222

[99] Phung, T.N. *et al.* (2016) Determining the effect of natural selection on linked neutral divergence across species. *PLoS Genet.* 12, e1006199

[100] Roux, C. *et al.* (2016) Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biol.* 14, e2000234

[101] Avise, J.C. *et al.* (1987) Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* 18, 489–522

[102] Avise, J.C. (1998) The history and purview of phylogeography: a personal reflection. *Mol. Ecol.* 7, 371–379

[103] Arbogast, B.S. and Kenagy, G.J. (2001) Comparative phylogeography as an integrative approach to historical biogeography. *J. Biogeogr.* 28, 819–825

[104] Smith, B.T. *et al.* (2012) Evaluating the role of contracting and expanding rainforest in initiating cycles of speciation across the isthmus of panama. *Proc. Biol. Sci.* 279, 3520–3526

[105] Carstens, B.C. *et al.* (2005) Investigating the evolutionary history of the pacific northwest mesic forest ecosystem: Hypothesis testing within a comparative phylogeographic framework. *Evolution* 59, 1639–1652

[106] Hickerson, M.J. and Cunningham, C.W. (2005) Contrasting quaternary histories in an ecologically divergent pair of low-dispersing intertidal fish (xiphister) revealed by multi-locus DNA analysis. *Evolution* 59, 344–360

[107] Hickerson, M.J. *et al.* (2006) Test for simultaneous divergence using approximate bayesian computation. *Evolution* 60, 2435–2453

[108] Carstens, B.C. *et al.* (2016) Community trees: Identifying codiversification in the páramo dipteran community. *Evolution* 70, 1080–1093

[109] Satler, J.D. and Carstens, B.C. (2016) Phylogeographic concordance factors quantify phylogeographic congruence among co-distributed species in the sarracenia alata pitcher plant system. *Evolution*

[110] Losos, J.B. (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecol. Lett.* 11, 995–1007

[111] Ricklefs, R.E. (2007) History and diversity: Explorations at the intersection of ecology and evolution. *Am. Nat.* 170, S56–S70

[112] Emerson, B.C. and Gillespie, R.G. (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends Ecol. Evol.*

[113] Vellend, M. (2005) Species diversity and genetic diversity: parallel processes and correlated patterns. *Am. Nat.* 166, 199–215

[114] Vellend, M. *et al.* (2014) Drawing ecological inferences from coincident patterns of population- and community-level biodiversity. *Mol. Ecol.* 23, 2890–2901

[115] Venkataraman, A. *et al.* (2015) Application of a neutral community model to assess structuring of the human lung microbiome. *MBio* 6

[116] Hubbell, S.P. (2005) The neutral theory of biodiversity and biogeography and stephen jay gould. *Paleobiology* 31, 122–132

[117] Etienne, R.S. and Olff, H. (2004) A novel genealogical approach to neutral biodiversity theory. *Ecol. Lett.* 7, 170–175

[118] Jabot, F. and Chave, J. (2009) Inferring the parameters of the neutral theory of biodiversity using phylogenetic information and implications for tropical forests. *Ecol. Lett.* 12, 239–248

[119] Burbrink, F.T. *et al.* (2015) Predicting community structure in snakes on eastern nearctic islands using ecological neutral theory and phylogenetic methods. *Proceedings of the Royal Society B: Biological Sciences* 282, 20151700

[120] Davies, T.J. *et al.* (2011) NEUTRAL BIODIVERSITY THEORY CAN EXPLAIN THE IMBALANCE OF PHYLOGENETIC TREES BUT NOT THE TEMPO OF THEIR DIVERSIFICATION. *Evolution* 65, 1841–1850

[121] Rosindell, J. *et al.* (2010) Protracted speciation revitalizes the neutral theory of biodiversity. *Ecol. Lett.* 13, 716–727

[122] Olszewski, T.D. and Erwin, D.H. (2004) Dynamic response of permian brachiopod communities to long-term environmental change. *Nature* 428, 738–741

[123] Wagner, P.J. *et al.* (2006) Abundance distributions imply elevated complexity of post-paleozoic marine ecosystems. *Science* 314, 1289–1292

[124] Kendall, D.G. (1948) On the generalized "Birth-and-Death" process. *Ann. Math. Stat.* 19, 1–15

[125] Kendall, D.G. (1948) On some modes of population growth leading to r. a. fisher's logarithmic series distribution. *Biometrika* 35, 6–15

[126] Engen, S. and Lande, R. (1996) Population dynamic models generating the lognormal species abundance distribution. *Math. Biosci.* 132, 169–183

[127] Engen, S. and Lande, R. (1996) Population dynamic models generating species abundance distributions of the gamma type. *J. Theor. Biol.* 178, 325–331

[128] McGill, B. (2003) Strong and weak tests of macroecological theory. *Oikos* 102, 679–685

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

[130] Etienne, R.S. and Haegeman, B. (2012) A conceptual and statistical framework for adaptive radiations with a key role for diversity dependence. *Am. Nat.* 180, E75–E89

[131] Rabosky, D.L. (2013) Diversity-dependence, ecological speciation, and the role of competition in macroevolution. *Annu. Rev. Ecol. Evol. Syst.* 44

[132] Rabosky, D.L. and Lovette, I.J. (2008) Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? *Evolution* 62, 1866–1875

[133] Adler, P.B. *et al.* (2007) A niche for neutrality. *Ecol. Lett.* 10, 95–104

[134] HilleRisLambers, J. *et al.* (2012) Rethinking community assembly through the lens of coexistence theory. *Annual Review of*

[135] Levine, J.M. and HilleRisLambers, J. (2009) The importance of niches for the maintenance of species diversity. *Nature* 461, 254–257

[136] Mayfield, M.M. and Levine, J.M. (2010) Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecol. Lett.* 13, 1085–1093

[137] Godoy, O. *et al.* (2014) Phylogenetic relatedness and the determinants of competitive outcomes. *Ecol. Lett.* 17, 836–844

[138] Kirchner, J.W. and Weil, A. (1998) No fractals in fossil extinction statistics. *Nature* 395, 337–338