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

Corresponding author: Rominger, A.J. (ajrominger@gmail.com).

Abstract

Whether or not biodiversity dynamics are governed by stable equilibria remains an unsolved question in ecology and evolution with important implications for our understanding of diversity and its conservation. Phylo/population genetic models and macroecological theory represent two primary lenses through which we view biodiversity. While phylo/population genetics provide an averaged view of changes in demography and diversity over timescales of generations to geological epochs, macroecology provides an ahistorical description of commonness and rarity across co-occurring species. Our goal is to combine these two approaches to gain novel insights into the non-equilibrium nature of biodiversity. We help guide near future research with a call for bioinformatic advances and an outline of quantitative predictions made possible by our approach.

1 Non-equilibrium, inference, and theory in ecology and evolution

The idea of an ecological and evolutionary equilibrium has pervaded studies of biodiversity from geological to ecological, and from global to local [1–7]. The consequences of non-equilibrium dynamics for biodiversity are not well understood and the need to understand them comes at a critical time when anthropogenic pressures are forcing biodiversity into states of rapid transition [8, 9]. The extent to which ecosystems are governed by non-equilibrial processes has profound implications for conservation, which are only just beginning to be explored [10].

Biodiversity theories based on assumptions of equilibrium, both mechanistic [4, 6, 7] and statistical [see the Glossary; 5, 11] have found success in predicting ahistorical patterns of diversity such as the species abundance distribution [4, 5, 12] and the species area relationship [4, 5]. These theories assume a macroscopic equilibrium in terms of these coarse-grained metrics, as opposed to focusing on details of species identity [such as in 9], although macroscopic and microscopic approaches are not mutually exclusive. Nonetheless, the equilibrium assumed by these theories is not realistic [13], and many processes, equilibrial and otherwise, can generate the same macroscopic, ahistorical predictions [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 non-equilibrial processes. The dynamic natures of evolutionary innovation and landscape change suggest that ecological theory could be greatly enriched by synthesizing its insights with inference from population genetic theory that explicitly accounts for history. This would remedy two shortfalls of equilibrial theory: 1) if theory fits observed ahistorical patterns but the implicit dynamical assumptions were wrong, we would make the wrong conclusion about the equilibrium of the system; 2) if theories do not fit the data we cannot know why unless we have a perspective

2

on the temporal dynamics underlying the generation of those data.

No efforts to date have tackled these challenges. We propose that combining insights from 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 ecological assemblages. The time is ripe to fully harness the vast amount of genetic and genomic data being generated at unprecedented scales [15–27] to address fundamental questions in ecology and evolution. However, we need a tool set of bioinformatic methods (Box 2) and meaningful predictions (section 6) grounded in theory to make use of those data. Data will take the form of both standard ecological metrics such as species abundances, as well as summaries of demographic/diversity dynamics inferred from genetics. Theory-based predictions will consist of connecting deviations from ecological theory and regions of parameter space with the dynamic processes inferred from genetics, all aided by new bioinformatic advances. We present the foundation of this tool set here.

2 Ecological theories and non-equilibrium

Neutral and statistical theories in ecology focus on macroscopic patterns, and the equilibrium presumed to be relevant to those patterns, but not the finer-grained properties of ecosystems. Non-neutral and non-statistical models [e.g., 6, 7] also invoke ideas of equilibrium in their derivation. However, these equilibria are defined at the level of individual populations, only accidentally leading to stable macroscopic patterns, and therefore do not fall within our primary focus. Here, we focus explicitly on simple yet predictive theories for their utility as null models, not because of a presumption of their realism.

Our goal throughout is not to validate neutral or statistical theories—quite the opposite, we propose new data dimensions, namely genetics, to help better test alternative hypotheses against these null theories, thereby gaining insight into what non-neutral and non-statistical

3

₆₂ mechanisms are at play in systems of interest.

2.1 Mechanistically neutral theory

₆₄ The neutral theory of biodiversity [NTB; 4] assumes that one mechanism—demographic
₆₅ drift—drives community assembly. By presuming that populations ~~do not differ in fitness nor~~
₆₆ ~~in~~ are equivalent in regards to competition or resource use, neutrality avoids the intractabil-
₆₇ ity of over parameterization and arrives as an equilibrium prediction when homogeneous
₆₈ stochastic processes of birth, death, speciation and immigration have reached stationarity.
₆₉ Thus neutrality in ecology is analogous to neutral drift in population genetics [4].

2.2 Statistical theory

₇₁ Rather than assuming any one mechanism dominates the assembly of populations into a
₇₂ community, statistical theories assume all mechanisms could be valid, but their unique in-
₇₃ fluence has been lost to the enormity of the system and thus the outcome of assembly is a
₇₄ community in statistical equilibrium [5, 11]. Thus these statistical theories are consistent
₇₅ with niche-based equilibria [28, 29] if complicated, individual or population level models were
₇₆ to be upscaled to entire communities. Thus strong and directional (i.e. non-equilibrial) niche
₇₇ and species interactions are potential drivers of non-equilibrium: such forces could prevent
₇₈ a system from attaining a neutral or statistical equilibrium.

₇₉ ~~Whatever mechanisms are at play, if they reach an equilibrium, statistical~~ Statistical
₈₀ theories derive predictions by condensing the many bits of mechanistic information down into
₈₁ ecological state variables [5] and then mathematically maximizing the entropy of predicted
₈₂ distributions ~~conditionally~~ conditional on those state variables. ~~One example is the~~ The
₈₃ maximum entropy theory of ecology (METE) ~~which assumes that the area of an ecosystem,~~
₈₄ ~~the total number of species in some taxonomic group, the total number of individuals in those~~
₈₅ ~~species, and the total metabolic rate of those individuals, capture all necessary information~~

4

to characterize a community [5]. While this theory finds widespread success in predicting ahistorical patternsof species abundance, body size, and spatial distribution is one such example [5]. METE can predict many ahistorical patterns, not just one [5], and while taxonomically and geographically widespread testing supports the theory [5, 12, 30] at single snapshots in time, it fails to match observed patterns in disturbed and rapidly evolving communities [5, 31].

2.3 Testing theory with ahistorical data

The vast majority of tests of neutral and/or statistical ecological theory have compared theoretical predictions to ahistorical data, most commonly the species abundance distribution [4, 5, 12], although other metrics concerning body size, trophic links, and spatial distributions have been used [e.g., 30][e.g., 5, 30, 31]. To test them these theoretical predictions, we need a robust measure of goodness of fit. The emerging consensus is that likelihood-based test statistics should be preferred [32]. The "exact test" of Etienne [33] has been extended by Rominger and Merow [34] into a simple z-score which can parsimoniously describe the goodness of fit between theory and pattern. We advocate its use in our proposed framework.

3 Inferring non-equilibrium dynamics

Unlocking insight into the dynamics underlying community assembly will help us overcome the limitations of analyzing ahistorical patterns with equilibrial theory. While the fossil record can elucidate deep time patterns for select, it is limited to groups [35], locales, and times [36], and therefore we require an approach that is applicable across taxa, space, and time. Populationcould be used for this task, it has limited temporal, spatial, and taxonomic resolution. Confronting these challenges in merging macroecology with paleontology is exciting and underway []. Here we instead focus on population/phylogenetics provide insights

~~on~~ phylogenetic insights into rates of change of populations and species ~~by processes of~~ because of the detailed characterization of demographic fluctuations, immigration, selection, ~~speciation, and extinction~~and speciation they provide. Bridging ecological theory with models from population/phylogenetics has great potential ~~to generate general-use tools to gain insight into the dynamics underlying contemporary biodiversity patterns~~ [37–46] ~~. However, bridging of these approaches~~ that has yet to ~~quantitatively~~ be fully realized. How we can best link the inferences of change through time from phylo/population genetics with inferences ~~of macroscopic ahistorical patterns from macroecology . If we can overcome this shortfall, it would be possible to finally test which biodiversity patterns we see today results from equilibrium or non-equilibrium~~from macroecology depends on what specific insights we can glean from genetic perspectives on demography and diversification.

One of the fundamental ~~tool~~tools allowing for complex historical inference with population genetic data is coalescent theory [47–51]. ~~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.~~ Coalescent theory allows for model-based estimation of historical parameters such as historical population size fluctuations [52, 53], divergence and/or colonization times [54, 55], migration rates [56], selection [57–59], and complex patterns of historical population structure [60, 61] and gene flow [62, 63]. This approach can also be put in a multi-species, community context ~~thanks to advances in historical~~ via hierarchical demographic models [64–68], even when only small numbers of genetic loci are sampled from populations [69].

These modeled demographic deviations from neutral demographic equilibrium ~~can also~~ could potentially be condensed to ~~summary statistics, as in the case with ecological theory.~~ multi-species summary statistics similar to how Tajima's D detects selection or demographic expansion in a single species. In this case it measures the standardized difference between two theoretical derivations of genetic diversity that should be equal under neutral equilibrium, but different when non-equilibrium demography dominates [70–77].

4 Current efforts to integrate evolution into ecological theory

While quantitatively integrating theory from ecology, population genetics, and phylogenetics has not yet been achieved, efforts so far to synthesize perspectives from evolution and ecology have been hindered by one or more general issues: 1) lack of a solid theoretical foundation, 2) inability to distinguish multiple competing alternative hypotheses, 3) lack of comprehensive genetic/genomic data, 4) lack of bioinformatic approaches to resolve species and their abundances.

Community phylogenetics [37] was founded to understand the roles of competition and environmental filtering on community assembly and assumes key ecologically-relevant traits are conserved along phylogenies;. It lacks a solid theory on trait-mediated competition and recruitment and trait evolution [78]. While the field has progressed [e.g., 79], not enough attention has been given to using phylogenetic information to understand the historical contingencies at play in community assembly [80, 81].

The recent growth in joint studies of genetic and species diversity [43, 45, 82, 83] have been useful in linking population genetics with ecological and biogeographic theory. These joint studies could be bolstered by developing full joint models that link community assembly, historical demography and coalescent-based population genetics combined with next generation sequencing based community analysis approaches (see Box 1).

Phylogeographic studies of past climate change have provided insights of how specific groups have responded in non-equilibrium ways to perturbations [68, 84–86], 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 [87] have used the exact same

kind of genetic/genomic data we propose using to test ecological theory in any group, from microbes to macrobes. However, such studies to date 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).

4.1 Emerging approaches

Several approaches have been taken that better ground synthesis of ecological and evolutionary dynamics in theory by making explicit use of quantitative predictions from ecological theory, along with incorporating evolutionary dynamics. They fall into two groups: 1) improving the agreement between data and ecological theory using evolutionary information; and 2) inferring non-neutral or non-statistical mechanics by combining theory with history.

Jabot and Chave [88] used approximate Bayesian computation (ABC) 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 [89]. While these works both 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. Therefore, 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 across evolutionary snapshots of community assembly with the goal of understanding how changing evolutionary dynamics drive community assembly [90, 91]. Rominger et al. [31] 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. These evolutionary snapshot studies lack a quantitative reconciliation of mechanisms inferred by analyses of ahistorical theory with independently inferred dynamics.

5   What is needed now

A key limitation of using ahistorical theory to infer dynamic mechanisms is that multiple mechanisms can map onto the same ahistorical pattern [92–96]. This means that even when a theory describes the data well, we do not know the dynamics that led to a good fit.

Quantitative theoretical foundations and direct information about dynamics can break this many-to-one mapping of mechanism onto theoretical prediction. This nicely parallels with calls to incorporate additional information into community ecology and macroecological studies [97]. Here we propose a much needed framework for integrating the dynamics inferred from population and phylogenetic approaches with ahistorical, equilibrial ecological theory. There are two complementary options for approaching our method:

- 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, while also fitting population genetic and/or phylogenetic models to genetic data 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.

5.1   What we could gain from this framework

Using our suggested framework, 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 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[98–100], but lacks a link this back to ecological assembly mechanisms. 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).

6   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

10

well as information on some portion of their genomes such that we can estimate historical demography and diversification. In Box 2 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~~community-level population genetics. 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.

6.1 Cycles of non-equilibrium

Ecosystems experience regular disturbances which can occur on ecological time-scales, such as primary succession, or evolutionary time scales, such as evolution of novel innovations that lead to new ecosystem processes [101, 102]. We hypothesize that these regular disturbances can lead to cycles of non-equilibrium in observed biodiversity patterns.

Figure 1 derives from comparing summaries of deviation from neutral/statistical equilibrium on the y-axis and deviations from equilibrial demography/diversification on the x-axis. Trajectories of biodiversity assemblages through this space shows how we hypothesize biodiversity to transition between different phases of equilibrium and non-equilibrium. Deviations from ecological theory can be quantified by the previously discussed exact tests [33] and z-scores [34], while many statistics are available to quantify departure from demographic/diversification steady state including the previously discussed Tajima's D. 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

11

- 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~~ models 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 ~~,~~ or be jointly observed, while ahistorical ecological theories have been fit to those abundance ~~data~~ proxies, and models of population demography and/or diversification have

12

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.

6.2   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 Box 1). Taken as a whole, systems in which ecological metrics deviate from equilibrial theory while demographic and macroevolutionary 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 [e.g., 5].

13

6.3 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 environment—then the system is open to evolutionary innovation. Such innovation could take the form of elevated speciation or long-distance immigration, relating to the idea that community assembly is a race between processes with potentially different, but stochastic rates [43]. Speciation and sweepstakes immigration/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).

6.4 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 (Fig. 1 Panel I to IV to III). Evolutionary change would have to be extremely rapid to force ecological metrics out of equilibrium, thus we would expect to see phylogenetic signals of adaptive radiation, and corresponding signals of strong selection in genomic-scale sequence data.

14

6.5   Ecological and evolutionary relaxation

Ecological metrics can return to equilibrium either by ecological means or by evolutionary means. 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. Given sufficient time in ecological equilibrium, this time averaged demographic record revealed by genetic inference will likely also re-equilibrate.

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

Inference of community dynamics needs to expand the dimensions of data it uses to draw conclusions about assembly of biodiversity and whether/when this is an equilibrial or non-equilibrial process. The need for more data dimensions is supported by others [14]; however, accounting for the complexities of history by explicitly linking theories of community assembly with theories of evolutionary genetics is novel, and made possible by advances in:

1. high throughput sequencing (Box 1) that allow genetic samples to be economically and time-effectively produced on unprecedented scales

2. bioinformatic methods (Box 2) that allow us to make sense of these massive community-wide genetic/genomic datasets

3. theory development (section 6) that provides meaningful predictions to test our new bioinformatic approaches

This approach is a fertile cross pollination of two fields that, while successful in their

own right, are enhanced by their integration. While comparative historical demographic models are advancing [64–68], testing community-scale hypotheses with multi-taxa data would be profoundly improved and enriched if population genetic model were grounded in macroecological and biogeographic theory. What is more, it has been long recognized that models in community ecology have been overly reliant on ahistorical patterns, such as the species abundance distribution, which are by themselves often insufficient for distinguishing competing models of assembly [14]. The field is ready to fully merge these two approaches using the wet lab, bioinformatic, and theoretical-conceptual approaches we have promoted here . This time to do so is now, as scientists face an increasingly non-equilibrium world and its consequences for our fundamental understanding of what forces govern the diversity of life and how we can best harmonize human activities with it.

Acknowledgements

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

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

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

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

[7] 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

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

[9] Blonder, B. *et al.* (2015) Linking environmental filtering and disequilibrium to biogeography with a community climate framework. *Ecology* 96, 972–985

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

17

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

[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] 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

[15] Yu, D.W. *et al.* (2012) Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution* 3, 613–623

[16] Pompanon, F. *et al.* (2012) Who is eating what: diet assessment using next generation sequencing. *Molecular ecology* 21, 1931–1950

[17] Taberlet, P. *et al.* (2012) Towards next-generation biodiversity assessment using dna metabarcoding. *Molecular Ecology* 21, 2045–2050

[18] Ji, Y. *et al.* (2013) Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology letters* 16, 1245–1257

[19] Zhou, X. *et al.* (2013) Ultra-deep sequencing enables high-fidelity recovery of biodiversity for bulk arthropod samples without pcr amplification. *Gigascience* 2, 4

[20] Tang, M. *et al.* (2014) Multiplex sequencing of pooled mitochondrial genomes—a crucial step toward biodiversity analysis using mito-metagenomics. *Nucleic acids research* 42, e166–e166

[21] Bohmann, K. *et al.* (2014) Environmental dna for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution* 29, 358–367

18

[22] Gibson, J. *et al.* (2014) Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through dna metasystematics. *Proceedings of the National Academy of Sciences* 111, 8007–8012

[23] Shokralla, S. *et al.* (2015) Massively parallel multiplex dna sequencing for specimen identification using an illumina miseq platform. *Scientific reports* 5, 9687

[24] Linard, B. *et al.* (2015) Metagenome skimming of insect specimen pools: potential for comparative genomics. *Genome biology and evolution* 7, 1474–1489

[25] Leray, M. and Knowlton, N. (2015) Dna barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. *Proceedings of the National Academy of Sciences* 112, 2076–2081

[26] Dodsworth, S. (2015) Genome skimming for next-generation biodiversity analysis. *Trends in plant science* 20, 525–527

[27] Liu, S. *et al.* (2016) Mitochondrial capture enriches mito-dna 100 fold, enabling pcr-free mitogenomics biodiversity analysis. *Molecular ecology resources* 16, 470–479

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

[29] Neill, C. *et al.* (2009) A competitive coexistence principle? *Oikos* 118, 1570–1578

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

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

[32] Baldridge, E. *et al.* (2016) An extensive comparison of species-abundance distribution models. *PeerJ* 4, e2823

[33] Etienne, R.S. (2007) A neutral sampling formula for multiple samples and an 'exact' test of neutrality. *Ecology letters* 10, 608–618

[34] Rominger, A.J. and Merow, C. (2017) meter: an r package for testing the maximum entropy theory of ecology. *Methods in Ecology and Evolution* 8, 241–247

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

[36] 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

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

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

[39] 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

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

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

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

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

454   [44] Vellend, M. and Geber, M.A. (2005) Connections between species diversity and genetic
455   diversity. *Ecol. Lett.* 8, 767–781

456   [45] Papadopoulou, A. *et al.* (2011) Testing the Species–Genetic diversity correlation in the
457   aegean archipelago: Toward a Haplotype-Based macroecology? *Am. Nat.* 178, 241–255

458   [46] Dexter, K.G. *et al.* (2012) Historical effects on beta diversity and community assembly
459   in amazonian trees. *Proc. Natl. Acad. Sci. U. S. A.* 109, 7787–7792

460   [47] Hudson, R.R. (1983) Properties of a neutral model with intragenic recombination.
461   *Theor. Popul. Biol.* 23, 183–201

462   [48] Tajima, F. (1983) Evolutionary relationship of DNA sequences in finite populations.
463   *Genetics* 105, 437–460

464   [49] Kingman, J.F.C. (1982) The coalescent. *Stochastic Process. Appl.* 13, 235–248

465   [50] Kingman, J.F.C. (1982) On the genealogy of large populations. *J. Appl. Probab.* 19a,
466   27–43

467   [51] Rosenberg, N.A. and Nordborg, M. (2002) Genealogical trees, coalescent theory and
468   the analysis of genetic polymorphisms. *Nat. Rev. Genet.* 3, 380–390

469   [52] Kuhner, M.K. *et al.* (1998) Maximum likelihood estimation of population growth rates
470   based on the coalescent. *Genetics* 149, 429–434

471   [53] Slatkin, M. and Hudson, R.R. (1991) Pairwise comparisons of mitochondrial DNA
472   sequences in stable and exponentially growing populations. *Genetics* 129, 555–562

473   [54] Charlesworth, D. (2010) Don't forget the ancestral polymorphisms. *Heredity* 105,
474   509–510

[55] 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

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

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

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

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

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

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

[62] 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

[63] 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

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

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

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

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

[68] 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*

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

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

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

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

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

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

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

[76] 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

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

[78] 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

[79] Sukumaran, J. *et al.* (2016) Machine learning biogeographic processes from biotic patterns: A new trait-dependent dispersal and diversification model with model choice by simulation-trained discriminant analysis. *Systematic biology* 65, 525–545

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

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

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

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

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

[85] 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

[86] 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

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

[88] 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

[89] 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

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

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

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

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

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

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

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

[97] 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

[98] 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

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

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

[101] Van Valen, L. (1973) A new evolutionary law. *Evolutionary theory* 1, 1–30

[102] Erwin, D.H. (2008) Macroevolution of ecosystem engineering, niche construction and diversity. *Trends in Ecology & Evolution* 23, 304–310

[103] Krehenwinkel, H. *et al.* (2016) A cost-efficient and simple protocol to enrich prey dna from extractions of predatory arthropods for large-scale gut content analysis by illumina sequencing. *Methods in Ecology and Evolution*

[104] Elbrecht, V. and Leese, F. (2015) Can dna-based ecosystem assessments quantify species abundance? testing primer bias and biomass—sequence relationships with an innovative metabarcoding protocol. *PloS one* 10, e0130324

[105] Coissac, E. *et al.* (2016) From barcodes to genomes: extending the concept of dna barcoding. *Molecular ecology*

[106] Nielsen, H.B. *et al.* (2014) Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nature biotechnology* 32, 822–828

[107] Angly, F.E. *et al.* (2014) Copyrighter: a rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. *Microbiome* 2, 11

[108] Saitoh, S. *et al.* (2016) A quantitative protocol for dna metabarcoding of springtails (collembola) 1. *Genome* 59, 705–723

[109] Overcast, I.A. and Hickerson, M.J. (in prep.) Integrating community assembly models and comparative population genetics.

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

[111] Kelleher, J. *et al.* (2016) Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLoS Comput Biol* 12, e1004842

Boxes

Box 1: Wetlab techniques

Next generation sequencing (NGS) technology has ushered in a revolution in evolutionary biology and ecology. Whole genome and transcriptome sequences can now be generated even for non-model organisms (Ellegren 2014) and tools like RAD sequencing (Hohenlohe et al. 2015) or ultraconserved elements (Faircloth et al. 2012) allow researchers to study intra - and interspecific genetic variation of thousands of loci in parallel. This progress led to fascinating new insights into the structure of the tree of life (Misof et al. 2014) and the genetic basis of adaptation and speciation (Jones et al. 2012; Soria-Carrasco et al. 2014). Current phylogenomic and population genomic studies are usually based on large numbers of loci for a limited number of taxa or populations. This tactic is reversed in NGS based community analyses, which allow researchers to characterize the composition of complex species communities based on a limited set of informative barcode markers. The large scale recovery from bulk samples (e.g. passive arthropod traps) of species richness, food web structure, and cryptic species promise unprecedented new insights into ecosystem function and assembly [17, 22, 23, 103]. Two approaches, differing in cost and effectiveness, have emerged.

Metabarcoding describes the targeted PCR amplification and next generation sequencing of short DNA barcode markers (typically 300-500 bp) from community samples [15, 18]. The resulting amplicon sequences can be clustered into OTUs or grafted onto more well supported phylogenies. Even minute traces of taxa in environmental samples can be detected using metabarcoding [21]. Amplicon sequencing is cheap, requires a small workload, and thus allows rapid inventories of species composition and species interactions in whole ecosystems [16, 22, 25]. However, the preferential amplification of some taxa during PCR leads to highly skewed abundance estimates [15, 104] from metabarcoding libraries.

Metagenomic approaches, in contrast, avoid marker specific amplification bias by sequencing libraries constructed either from untreated genomic DNA [20, 24, 26], or after targeted enrichment of genomic regions [27]. While being more laborious, expensive and computationally demanding than metabarcoding, metagenomics thus offers improved accuracy in detecting species composition and abundance [19]. 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 [105]. 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 DNA clusters are popular targets [26, 105]. In contrast, microbial metagenomic studies ~~now~~ routinely assemble complete genomes and characterize gene content and metabolic pathways even from complex communities [106]. This allows unprecedented insights into functional genetic process underlying community assembly and evolutionary change of communities to environmental stress. Such whole genome based community analysis is not yet feasible for macroorganisms. However, considering the ever increasing throughput and read length of next generation sequencing technology, as well as growing number of whole genomes, it might well become a possibility in the near future, opening up unprecedented new research avenues for community ecology and evolution.

Box 2: Bioinformatic advances

While species richness can be routinely identified by sequencing bulk samples using high throughput methods, estimating species abundance remains challenging [104] 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 high throughput data. 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 model.

Model-free abundance estimation. We propose a pipeline (Fig. I) where raw reads are generated and assembled into a phylogeny using standard approaches, and potentially aided by additionally available sequence data in a super tree or super matrix approach. The numbers of sequences assigned to each terminal tip are then used in a Bayesian hierarchical model which seeks to estimate the true number of organisms representing each terminal tip, accounting for sequencing biases originating from, e.g. primer affinity and copy number differences between taxa. Information on phylogenetic relatedness can inform modeled correlations in biases between taxa [107, e.g. copy number is known to be phylogenetically conserved at least in microbes]. This approach is particularly tailored to metabarcoding data. In a potentially powerful extension, and thanks to the proposed Bayesian framework, information from sequencing experiments that seek to calibrate metabarcoding studies [e.g., 103, 108] can be used to build meaningfully informative priors and improve model accuracy. Through a simulation study (described in the supplement) we show that true underlying abundances can be accurately estimated (Fig. II).

Joint inference of community assembly and population genetic models. Coupling individual-based, forward-time models of community assembly with backwards-time hierarchical multi-taxa coalescent models permits inference about the values of the parameters in both models. This framework is flexible enough to incorporate multiple refugia, colonization routes, ongoing migration and both neutral and deterministic processes of assembly on time scales of hundreds of thousands of years (Fig. III). A forthcoming implementation [gimmeSADπ; 109] ~~simulates an~~ jointly models a forward-time individual-based ~~forward time community dispersal model [110] linked with~~ neutral community assembly process [110] and corresponding expectations of community level genetic diversity and divergence using the

msPrime coalescent simulator [111]. 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.

Glossary

ahistorical Patterns or theories which do not contain information about the historical processes that gave rise to them.

Approximate Bayesian computation (ABC) A method of calculating an approximate posterior sample of parameters in a complex model whose likelihood function cannot be analytically solved by simulating realizations of the model, computing summary statistics from those realizations, and probabilistically accepting or rejecting the parameter values leading to those summary statistics based on their agreement with the observed statistics computed from the real data.

coalescent A stochastic, backwards in time ~~model approach in population genetics in~~ population genetic model in which alleles in the sample are traced to their ancestors under demographic models of interest.

equilibrium

hierarchical model

statistical equilibrium In the context of biodiversity, a description of a steady state arrived at not by the force of one or a few deterministic mechanisms but by the stationary, statistical behavior of very large collections of mechanistic drivers acting on large assemblages of organisms.

Tajima's D A metric of non-stationary evolution computed as the difference between two distinct derivations of theoretical genetic diversity. If neutral molecular evolution in a constant population holds, both derivations should be equal, and otherwise if assumptions of constant demography and neutral selection are violated, will be unequal.

Outstanding Questions

1. How can fossil data be best integrated with ahistorical ecological theory and diversity dynamics as informed by phylogenetic and population genetic inference?

2. How can functional genomics be used to better distinguish between purely demographic and niche-based drivers of non-equilibrium? Can understanding the ~~functional~~ gene content of genomes, gene expression patterns or occurrence of mutations across taxa in a community help predict potential for non-equilibrium responses to future perturbations? ~~Functional genomics is still very much in development (see Box 1), but future prospects are exciting.~~

3. How can relative abundance data derived from ancient DNA and fossil data be leveraged within a joint model that generates predictions of spatiotemporal distributions of genetic polymorphism and species abundances? One such opportunity is the availability of highly resolved estimates of relative abundance distributions of forest tree assemblages that are derived from paleo-pollen data (Dawson et al. 2016) which could allow for joint inference in conjunction with assemblage-level genomic sampling. Likewise, obtaining community-level DNA preserved in lake sediments sampled at different

⁷²⁰ late Pleistocene and Holocene could provide for a whole new lense for testing models

⁷²¹ that account for historical dynamics at both evolutionary and ecological time scales

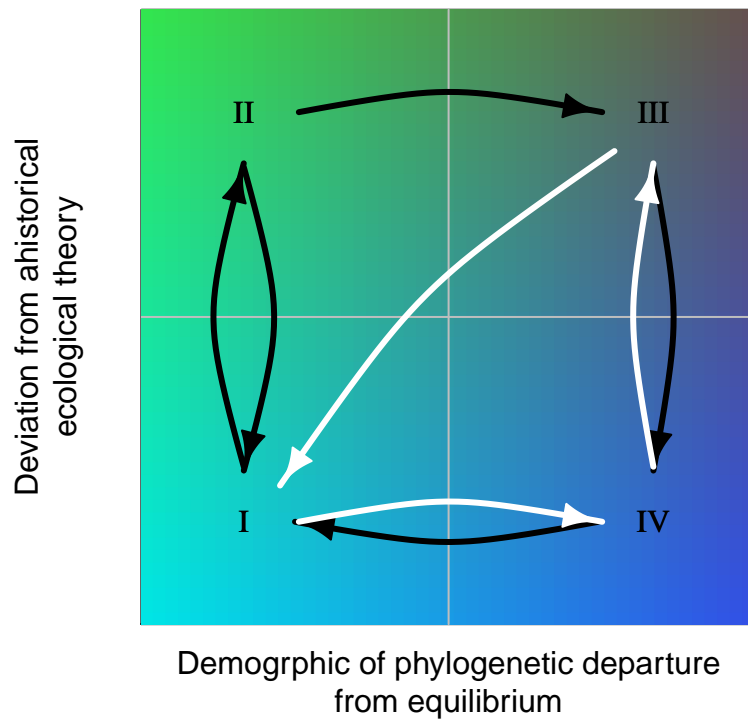⁷²² (Capo et al. 2016).

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~~text. 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.
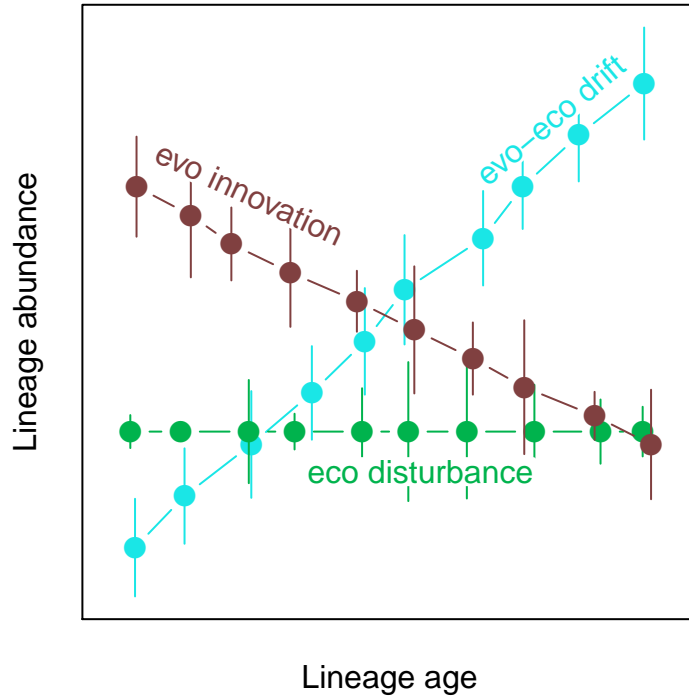
Figure 2: Hypothesized relationships between lineage age and abundance under different evo-ecological scenarios. Colors correspond to panels in Figure 1: teal is evo-ecological equilibrium; green is rapid transition to ecological non-equilibrium following short timescale disturbance; dark brown is non-equilibrium in both ecological and evolutionary metrics.
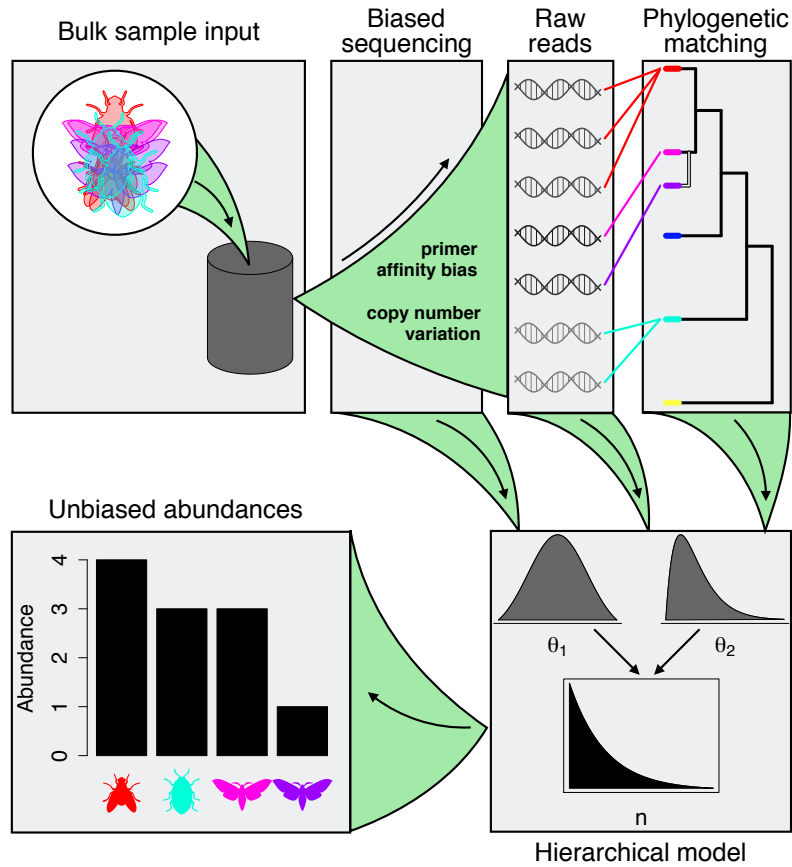
Box 2 figures

Figure I: Pipeline to estimate true abundances from metabarcoding data. The pipeline follows sequence generation, matching sequences to a phylogeny (generated from the sequences themselves, or better yet from higher coverage data) and finally Bayesian hierarchical modeling leading to abundance estimates.
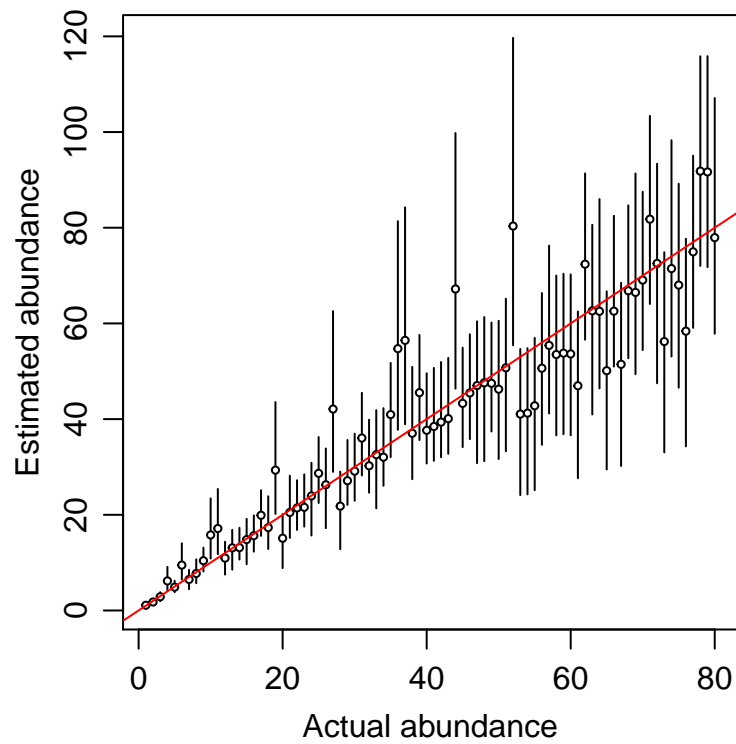
Figure II: Demonstration of agreement between actual and estimated abundances. Actual (simulated) abundances are on the x-axis, which the y-axis shows estimated abundances (error bars are 95% maximum credible intervals). The simulation study is described in the supplement.
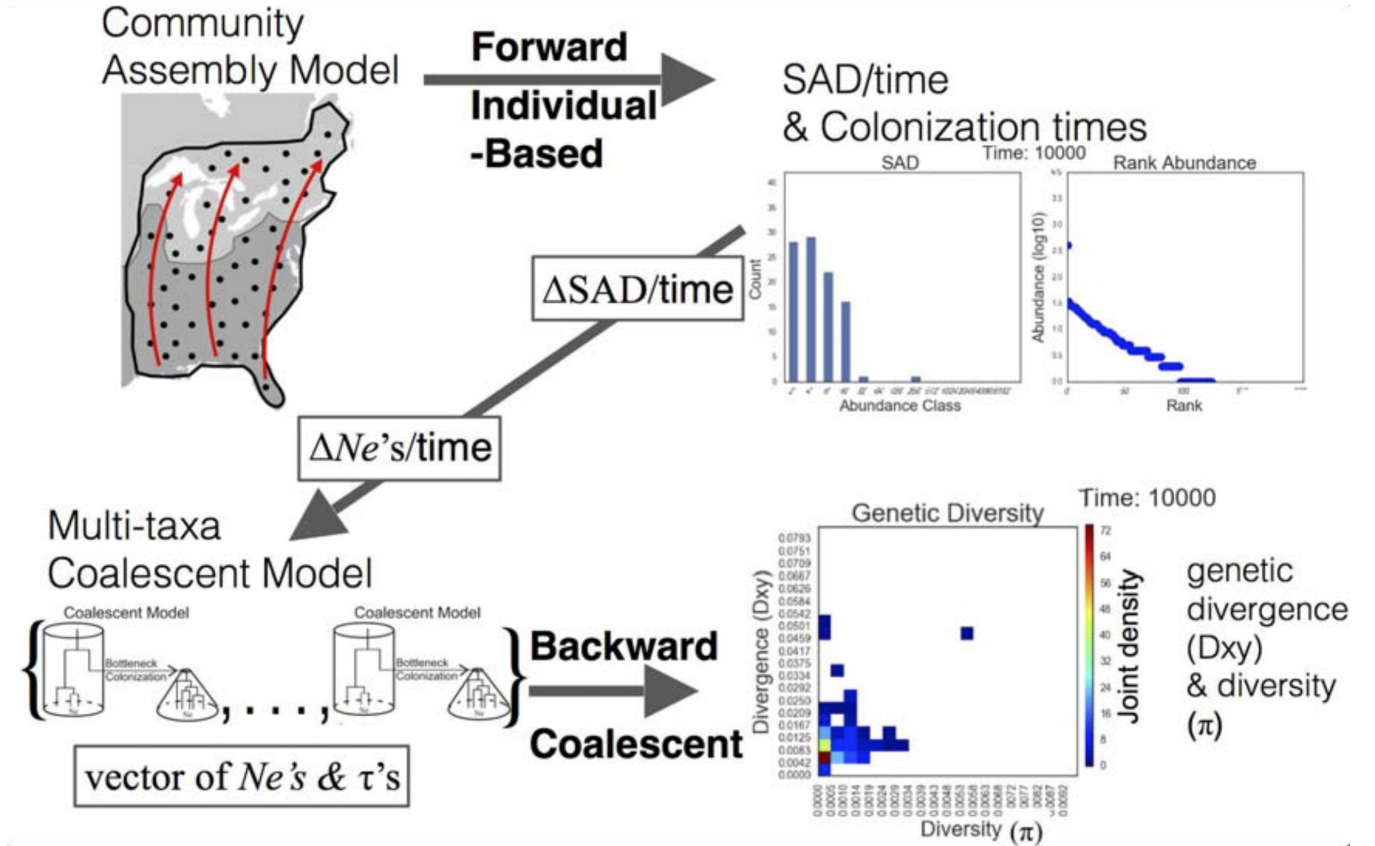
Figure III: The gimmeSADπ pipeline. The forward time models involves multi-regional expansion generating local abundance distributions over time with heterogeneity in colonization times. These temporally dynamic local abundances are re-scaled into local n_e distributions over time to generate multi-species genetic data the the coalescent, which is summarized here with a time-dependent joint spectrum of genetic diversity statistics.