Community assembly on isolated islands: Macroecology meets evolution

A. J. Rominger1\*, K. R. Goodman1\*, J. Y. Lim2\*, F. S. Valdovinos3\*, E. E. Armstrong1,4+, L. E. Becking1,5, G. M Bennett6, M. S. Brewer1,7+, D. D. Cotoras2,8+, C. P. Ewing4, J. Harte1, N.D. Martinez3, P. M. O’Grady1, D. M. Percy9, D. K. Price4, G. K. Roderick1, K. L. Shaw10, D. S. Gruner11#, R. G. Gillespie1#

1. Department of Environmental Science, Policy, and Management, University of California, Berkeley, California 94720-3114
2. Department of Integrative Biology, University of California, Berkeley, California 94720-3140
3. Pacific Ecoinformatics and Computational Ecology Lab, Berkeley, California 94703
4. Department of Biology, University of Hawaii, Hilo, Hawaii, 96720-4091
5. IMARES Wageningen UR, P.O. Box 57, 1780 AB Den Helder, The Netherlands
6. Department of Integrative Biology, University of Texas, Austin, Texas 78712
7. Department of Biology, East Carolina University, Greenville, NC 27858
8. Department of Entomology and Center for Comparative Genomics. California Academy of Sciences. San Francisco, CA 94118
9. Department of Life Sciences, Natural History Museum, London, UK SW7 5BD
10. Department of Neurobiology and Behavior, Cornell, Ithaca, New York 14853-7601
11. Department of Entomology, University of Maryland, College Park, Maryland 20742-4454

\*Contributed equally; # co-senior authors; +current affiliation

corresponding author: R. G. Gillespie, gillespie@berkeley.edu.

Keywords: networks, maximum entropy, arthropods, population genetics, chronosequence, Hawaii

Running title: Community assembly on isolated islands

Number of words in the abstract: XXX

Number of words in main body of the paper: XXXX

Number of references: XX

#### Aim

Understanding how ecological and evolutionary processes together determine biodiversity patterns remains a central goal in biology. Guided by ecological theory, we use data from multiple arthropod lineages endemic to the Hawaiian archipelago to explore the interplay between ecological (population dynamics, dispersal, trophic interactions) and evolutionary (genetic structuring, adaptation, speciation, extinction) processes. With this framework, our goal is to show how communities develop over ecological-evolutionary time, and the dynamic feedbacks involved in their assembly.

#### Location

The main Hawaiian Islands (19°-22°N, 155°-160°W).

#### Methods

To infer processes involved in early diversification we synthesize data on genetic population structure of select arthropod species across the Hawaiian Archipelago. Over the range of geological ages of the current high islands ( My to 5 My) we also develop and analyze a plant-herbivore bipartite interaction network. We compare the structure of these networks, measured by nestedness, modularity and the degree distribution, with theoretical predictions derived from the principle of maximum information entropy.

#### Results

Based on the perspective provided by the island chronosequence and population genetic information, we demonstrate that species at lower trophic levels develop population genetic structure at smaller temporal and spatial scales than species of higher trophic levels. Network nestedness decreased while modularity increased with habitat age. Single island endemics exhibit more specialization compared to more broadly distributed species. Plant-herbivore networks exhibit increasing deviation from theoretical predictions of food web structure across the increasing age gradient. Finally, at intermediate aged sites, communities showed the least deviation from maximum entropy theory of ecology (METE) prediction.

#### Main conclusions

We use available data from an oceanic archipelago to develop a framework to study the feedbacks between ecological and evolutionary processes in the construction of communities. Migration among localities is necessary for early community assembly on early successional sites. Populations of species at lower trophic levels develop genetic differentiation earlier in the chronosequence, compared to higher trophic levels. This suggests that *in situ* processes resulting in species differentiation have an impact on community assembly earlier in the chronosequence for herbivores than for predators. For insect herbivore networks, the pattern of decreasing nestedness and increasing modularity may also indicate a shift from assembly driven by immigration early on, to one based on *in situ* co-diversification with host plants as communities age. This possibility is supported by the observation that single island endemics (likely the product of *in situ* diversification) show more specialization at older sites compared to more broadly distributed species (those taxa more likely to be initial colonists). Finally, the observation that the younger and oldest aged sites, but not intermediate aged sites, deviate from the predictions of METE suggests that plant-herbivores communities only transiently achieve statistical state during the assembly process, presumably due to incomplete assembly from dispersal during early stages of assembly, and the increasing influence of island ontogeny on plant and herbivore populations on old islands. Future work will focus on extending the observations presented here with more detailed data and development of theory.

# Introduction

Contemporary biodiversity is a product of speciation, extinction and dispersal, contingent on the ecological interactions of organisms with the biotic and abiotic environment. The evolutionary history leading to any given ecological community must in some way shape contemporary ecological assemblages but because the processes of evolution and ecology occur on different temporal and spatial scales, disentangling the relative influence of local ecological mechanisms from historical evolutionary processes on biodiversity patterns remains a central challenge (Ricklefs, 2004).

Evolutionary processes of speciation and extinction are typically viewed as constraints on regional species pools, occurring in a manner largely removed from local ecology (Hubbell, 2001; Cavender-Bares *et al.*, 2009; Wiens, 2011), while ecological mechanisms conversely tend to be viewed as packing standing diversity into local communities through competition, facilitation, and neutral ecological drift (Hubbell, 2001; Tilman, 2004; Bascompte & Jordano, 2007; Borer *et al.*, 2014). While recent theoretical advances have provided greater insight on ecological drift (Hubbell, 2001; Rosindell & Phillimore, 2011), niche partitioning (Tilman, 2004), competition, predation (Borer *et al.*, 2014), and species interaction networks (Williams & Martinez, 2000; Brose *et al.*, 2006), these insights either do not contain realistic evolutionary assumptions (Ricklefs, 2006) or ignore it entirely. Other insights into evolutionary mechanisms have also emerged based on inference from contemporary patterns of species, genetic, or phylogenetic diversity (e.g., Wiens, 2011; Jetz *et al.*, 2012). However, the temporal dynamics of both the evolutionary mechanisms and their ecological consequences cannot be inferred from static contemporary patterns nor can we understand what constitutes meaningful change in a system without a baseline to compare to. Here we show how the effects on diversity of changing interactions across an ecological-evolutionary continuum can be identified by studying the biota of archipelagoes with islands sequentially arranged by age combined with predictions from idealized ecological theories (such as the unified neutral theory [Hubbell, 2001] or the maximum entropy theory of ecology [Harte 2011]). The different aged islands provide approximate snapshots of the eco-evolutionary process at different stages. Ecological theory provides an idealized “null” baseline against which to compare observed patterns.

## Hotspot oceanic archipelagos as model systems

Hotspot oceanic islands are opportune model systems for studying the interplay of local ecological mechanisms and evolutionary drivers of biodiversity patterns. Due to their sequential formation as the tectonic plate moves over a volcanic hotspot, such island systems offer a range of spatial and temporal scales over which to analyze the outcomes of ecological and evolutionary processes (Warren et al. 2015). While many archipelagos around the world share these biotic and geologic properties, the Hawaiian archipelago provides a particularly useful system for study because its linear geological chronology (Price & Clague, 2002), ecosystem developmental trajectories (Vitousek 2004), and phylogeographic patterns of biodiversity are each well characterized (Wagner & Funk, 1995). Moreover, studies of diversity across the islands have demonstrated patterns that are non-uniform across the island chronosequence with marked differences between lineages (e.g., Gruner 2007; Gillespie & Baldwin, 2009) that can be used to test for what biologically meaningful differences between lineages drive their disparate diversification patterns.

## Development of genetic structure

High levels of dispersal, and associated gene flow, among localities limit the extent to which populations can diverge genetically. However, when gene flow is low the ecologies of distinct populations in different localities are free to diverge through local selective pressures and drift, leading to diversification. Thus, the magnitude of genetic connectivity among populations provides a measure of the relative importance of dispersal-driven community assembly (dictated by processes removed from the local setting) in contrast to assembly by local (*in situ*) diversification in determining community composition. Using the chronosequence of the Hawaiian Archipelago, we can analyze populations from multiple sets of taxa across trophic guilds occurring in geological contexts from young to old. We predict that ecological processes will dominate in community assembly in young habitats with the importance of evolutionary processes increasing with habitat age. If evolutionary processes were not important, we would predict that communities would reach a statistical steady state through ecological processes alone. If, as we expect, evolutionary processes become increasingly important in community assembly over time, we would expect to observe associated deviations from an ecological null model of community assembly—provided by idealized ecological theory. Increase in population structure relative to other taxa or trophic groups could indicate whether sufficient time has passed along the chronosequence for the group of interest to experience significant evolutionary pressures.

## Macroecological metrics

Unified theories of biodiversity (e.g. Hubbell, 2001; Harte, 2011) by their nature provide a simplified view of ecology, but deviations from theory can provide insights as to which particular ecological patterns require additional biological mechanisms for their explanation (Harte, 2011). The maximum entropy theory of ecology (METE; Harte, 2011) provides predictions of species abundance distributions, species area relationships, and metabolic rate and network linkage distributions, for idealized ecological communities in which a system’s behavior is governed by a simple set of state variables. The principle of maximum information entropy (MaxEnt), from which METE is derived, is an established inference procedure that has yielded accurate predictions of diverse patterns in fields as varied as thermodynamics (Jaynes, 1957), economics (Golan *et al.*, 1996), forensics (Roussev, 2010), imaging technologies (Gull and Newton, 1986), and recently ecology (e.g. Phillips et al., 2006; Dewar and Porté, 2008; Harte, 2011). MaxEnt works by seeking the least-biased prediction of a distribution of interest (e.g. the distribution molecular velocities in the case of thermodynamics or of species abundances in the case of ecology) while constraining that prediction to be consistent with state variables describing the macroscopic attributes of the system (e.g. temperature or the total numbers of species and individuals). These predictions are the most ignorant possible predictions about the system. Studying the unique ecological conditions and evolutionary histories of real world systems that deviate from METE can provide insights into the processes driving ecological systems away from the statistical steady state it describes and toward alternate system states (Harte, 2011).

Ecological networks are a prime study focus because networks of interacting species embody both evolutionary processes such as coevolution (Bascompte & Jordano, 2007; Donatti *et al.*, 2011; Nuismer *et al.*, 2013; Thompson et al. 2013) and the ecology of trophic links. Thus they present an opportune starting place to study ecological and evolutionary feedbacks. The distribution of linkages in ecological networks can test whether plant-animal interaction networks assemble neutrally or through deterministic processes such as co-evolution of traits involved in foraging (Vázquez *et al.*, 2005). Analysis of other network metrics such as modularity (the degree to which species interact in semi-autonomous modules) and nestedness (the degree of asymmetry in interaction between specialists and generalists) can further illuminate underlying eco-evolutionary processes driving patterns of species interactions (Bascompte & Jordano, 2007; Donatti *et al.*, 2011; Nuismer *et al.*, 2013). In these networks, species with fewer interactions (i.e. more specialized species) will interact with a subset of the species with which generalists interact.

To gain insights into community assembly as it happens, we propose an integrative framework that harnesses advances in both evolutionary and ecological theory, placed in the context of age-structured archipelagos. Mechanistically simplified ecological theories such as neutral theory (Hubbell, 2001) or the maximum entropy theory of ecology (METE, Harte, 2011) can be used as powerful null models; deviations from expectations of the theories can flag biological phenomena warranting further study. Here we demonstrate how community-level data from age-structured island systems, combined with population genetic and phylogenetic data, can test the extent to which the evolutionary history behind such communities drives their deviation from theoretical expectations. We provide an initial test of this concept using a synthesis of published data on arthropod lineages in the Hawaiian Islands. We provide metrics of ecological and evolutionary dynamics across communities from settings that range in geological age from 500 yr to 5 myr. We estimate taxon-specific timelines for the development of population genetic structure for both herbivores and predators and couple these results with macroecological measures of community structure, using predictions from statistical steady state and ecological network theory to provide insights into changes in community structure over the extended timeframe provided by the island chronosequence.

# Methods

## Population genetic inference of structure among populations

To evaluate the balance between regional immigration and potential for local differentiation, we measured how molecular variation is partitioned among populations within species across locations of known substrate age on Hawaii Island and Maui. We compiled published and new data sets for a diversity of native Hawaiian arthropod groups that represent a spectrum of trophic levels (Table 1). New sequences were included for *Nesosydne* planthoppers (COI: data generated following the protocols in Goodman *et al.*, 2012 GenBank accession numbers XXX-XXX) and *Trioza* psyllids (COI, cytB: data generated following protocols in Percy 2003; GenBank accession numbers XXX-XXX). Samples were from the focal sites described above for the ecological analysis as well as from locations across Hawaii Island and Maui. These data are used to provide an estimate of how arthropod populations have accumulated genetic population structure within the focal sites of different geological age.

We used analysis of molecular variance (AMOVA) to examine how genetic variation is partitioned at two scales of population structure: among sites within volcanoes and among volcanoes on both Hawaii Island and the islands of the Maui Nui complex (Maui, Molokai, Lanai). All analyses of allozyme and DNA sequence data were performed in Arlequin v. 3.5 (Excoffier & Lischer, 2010) using the AMOVA procedure to compute , a measure of genetic variance, or, where possible,, an analog that incorporates genetic sequence information. The *Laupala* AFLP data were analyzed using TFPGA v. 1.3 (Miller, 1997), using the same hierarchical approach as described above. To provide a temporal framework for the population differentiation analysis we assembled divergence dating information from the literature for as many of the taxa as possible and additionally implemented a new divergence dating analysis for *Tetragnatha* spiders (see supplementary information).

To explicitly test the association between landscape age and potential for *in* situ genetic divergence we analyzed how within site varies with the geologic age of volcanoes on Hawaii Island and Maui Nui. For each volcano we calculated or (Excoffier & Lischer, 2010) for each taxon among sites within volcanoes.

## Construction and analysis of plant-herbivore networks

## To investigate how ecological patterns change as communities age, we selected four focal sites across the chronosequence and island ages (two on Hawaii Island, one on Maui and one on Kauai; Fig. 1). Focal sites were selected to have similar forest composition (dominated by *Metrosideros polymorpha*: Myrtaceae), elevation (1100-1400m), and rainfall (mean annual precipitation 2000-3000 mm).

We then constructed bipartite interaction networks between native herbivorous Hemiptera species and native plants at each of the study sites. Bipartite networks describe the topology of ecological interactions between two guilds of organisms (e.g. herbivores and their plant hosts). Quantitative information on the relative importance of interaction links can be incorporated into network analyses (Vázquez *et al.*, 2009); however, currently available data are restricted to binary networks, those that describe only the potential for interaction between any two species, but not the relative frequency of that interaction to each species.

We compiled species lists of all native herbivorous Hemiptera for each focal site from published species accounts (see supplement for full list). Species accounts and other published sources were used to determine the presence, probable presence, or probable absence of each species at each of our four focal sites. A documented presence was defined as a known specimen collected at the focal site; a probable presence was defined as a species whose abiotic tolerances and known geographic range (see supplement) overlap with a focal site but no known specimen exists confirming its presence. Probable absence was assumed when neither criteria for presence or probable presence are met. Two sets of species lists for each focal site were compiled: a conservative data set composed of only documented presence occurrences and a less conservative data set that also included probable presences.

Host plants for each species of Hemiptera were determined from published species accounts. Data on host plant use at each specific site were not available so we assumed that if a known host plant were present at a site it would eventually be used. Host plant occurrence in the focal sites was determined using distribution models for 1158 species of Hawaiian plants (Price, 2012). Each focal site was spatially joined in a geographic information system with all coincident plant distribution models that fell within its boundaries. Two sets of resulting focal site-specific networks were constructed: one using the conservative data set of Hemiptera species presences and the other using the less conservative data set.

We hypothesized that potential complex evolutionary feedbacks contributing to community assembly should result in departures from the predicted ecological statistical steady state . We used METE (Williams, 2010; Harte, 2011) to compute the statistical steady state for the distribution of the number of host plants used by each Hemiptera species (hereafter referred to as degree distribution). To evaluate how well METE predicts the data we simulated METE-conforming communities of the same number of species and links as observed. We then calculated the log-likelihood of each simulated data set and compared the resultant distribution of log-likelihoods under the hypothesis that METE is true, to the observed log-likelihood. This comparison is identical in approach to a z-score test using a Monte Carlo simulation to estimate the sampling distribution of log-likelihoods. R scripts (v3.1.1, R Core Team 2014) used for METE estimation and Monte Carlo methods are available in the supplement.

To investigate how speciation may in part drive network patterns and deviations from those predicted by idealized ecological theory, we analyzed the number of links assigned to each Hemiptera species (the degree distribution) separately for single island endemics (those species found on only one island and thus likely derived from *in situ* diversification) versus multi-island endemics (those species found on multiple islands). Although multiple processes can lead to a species being a single-island endemic (Whittaker et al. 2008), such taxa provide a proxy for how much speciation occurs within islands. To compare species’ degree distributions between single-island endemics and multi-island endemics across sites of different ages we conducted a generalized linear model with binomial error, treating site identity as a categorical predictor. Binomial errors effectively account for network size due to the bounded support of the binomial distribution.

To understand how other network properties change with ecosystem substrate age, we calculated two widely used descriptive network metrics across sites – nestedness and modularity. Nestedness describes the degree of asymmetry of species interactions connecting specialists and generalists (Bascompte & Jordano, 2007; Ulrich *et al.*, 2009). We calculated nestedness using the NODF metric (Almeida-Neto *et al.*, 2008) as implemented in the R package *vegan* (Oksanen *et al.*, 2013) and modularity using a variety of algorithms implemented in the R package *igraph* (Csardi & Nepusz, 2006). These metrics are not directly comparable across networks of different size and connectance (Ulrich *et al.*, 2009), so for each metric in each network we calculate z-scores using a null model that randomizes network structure while maintaining certain aggregate network properties (Ulrich *et al.*, 2009). These z-scores are calculated as the difference between the observed network metric minus the mean of the null model divided by the null model standard deviation, or (). Because z-scores can be highly sensitive to the choice of null model (Ulrich *et al.*, 2009) we implemented both a probabilistic null model (Bascompte & Jordano, 2007) and a null model that strictly constrains the degree distributions of plants and herbivores (Ulrich *et al.*, 2009). The probabilistic null using the frequency of interactions as the probability that a randomized link gets assigned to that cell in the interaction matrix (Bascompte & Jordano, 2007); thus the probabilistic null constrains row and column sums in probability but not absolutely.

# Results

## Population genetic inference of structure among populations

The analysis of molecular variance (AMOVA) revealed significant genetic population structure from the smallest to the largest spatial scales examined, all within a very recent timeframe. For mitochondrial loci, statistically significant molecular variation partitioned among sites within volcanoes ranged from 0.037–0.92 and among volcanoes from 0–0.30. Corresponding variation at multilocus nuclear loci among-sites within volcanoes ranged from 0.21–0.58 and among volcanoes, 0.04–0.34. Taxa in the lower trophic levels (herbivorous sap-feeding Hemiptera: planthoppers and psyllids) had as much or more molecular variation partitioned at the among-site, within volcano level than the among volcano level, while the predatory spiders were less structured at localities within volcanoes compared to among them (Table 1). The analysis of genetic population structure across the chronosequence of localities revealed a similar pattern. The herbivores show high genetic population structure among localities even on young volcanoes (Fig. 2). By contrast, predatory spiders exhibited higher genetic population structure between volcanoes, but not within sites on the same volcano.

The observed levels of genetic divergence have evolved rapidly. For example, for species from Hawaii Island for which phylogenetic data provide divergence times, estimates of dates of species divergence range from 0.34–1.15 million years, with additional within-species genetic divergence having developed subsequently. That some of these estimates are older than the known age of the Big Island suggests that genetic divergence pre-dates their colonization to Hawaii Island, or alternatively that estimates include sampling error. For the one species where population genetic data was used to estimate divergence times between populations, herbivorous *Nesosydne* planthoppers, it was determined that populations diverged as little as 2,600 years ago (Goodman *et al.*, 2012) (Table 1).

## Analysis of plant-herbivore networks

The degree distribution of Hemiptera species varied across the chronosequence; with both the youngest and oldest sites deviating most from the statistical steady state maximum entropy predictions (Fig. 3). In the intermediate-aged site of Kohala, deviations from maximum entropy are not significantly different from predictions of maximum entropy.

The generalized linear model revealed significant differences between the degree distributions of single-island endemics (species whose distributions are restricted to only one island) versus archipelagic endemics that are found across multiple islands (Fig. 3). Single-island endemics show significantly lower degree distributions overall (i.e., more specialization) compared to more generalist species found across multiple islands. Furthermore, single-island endemics use more host plant species on the intermediate-aged Maui site. The slightly younger Kohala shows increased generalization for both single island endemics and archipelago endemics.. However, when considering the degree distribution defined by trophic links to plant genera instead of plant species, the pattern of increased generalization holds for Kohala but endemics on Maui no longer show a difference in their degree distributions from other island endemics. This change in pattern suggests that increased generality of Maui endemics may be driven by increased plant species diversity within genera on that island.

Network nestedness decreased with habitat age while modularity increased (Fig. 4). This trend was recovered in networks constructed from both more and less stringent geographic criteria (supplemental Fig. S3). Choice of null model changed the magnitude of modularity and the sign of nestedness z-scores; however, the relative pattern of decreasing nestedness and increasing modularity remained across the different null models used to standardize network metrics (supplemental Fig. S2). The patterns are also robust to sampling intensity, as demonstrated by a rarefaction analysis (supplemental Fig. S4).

# Discussion

## Development of genetic population structure at different trophic levels

The analysis of available genetic data presented here indicate that divergence is occurring within the islands at small spatial scales and over short time periods (Table 1, Fig. 2). Furthermore, the scale of population structure varies with trophic position, with structure developing in sap-feeding herbivore lineages at smaller scales compared to detritivorous crickets and predatory spiders (Table 1, Fig. 2). Structure within species may allow populations to take independent evolutionary trajectories, especially when aided by other evolutionary processes acting differentially across species ranges. A variety of factors have been associated with the genetic divergence of populations and species in lineages described here, including combinations of genetic drift associated with geographic isolation (Percy, 2003; Mendelson & Shaw, 2005; O’Grady *et al.*, 2011; Goodman *et al.*, 2012), adaptation associated with ecological interactions of competition, predation, and mutualism (Gillespie, 2004; Roderick & Percy, 2008; Brewer *et al.* 2015), and sexual signaling (Mendelson & Shaw, 2005; Percy *et al*. 2006; Magnacca *et al.*, 2008; Goodman *et al.*, 2015).

The sap-feeding Hemiptera group of *Nesosydne* planthoppers (Goodman *et al.*, 2012) provides evidence that some period of geographic isolation preceded divergence of sexual signals (Goodman *et al.*, 2015). Shifts in plant host use are also associated with diversification in this group (Roderick & Percy, 2008). In a phylogenetic study of a radiation of sapfeeding *Nesophrosyne* (Cicadellidae) leafhoppers, host plant specialization was implicated in driving species radiations up until approximately 1 million years ago, consistent with the hypothesis that plant niches were mostly exhausted on Maui and speciation in the younger landscape of Hawaii Island shifted to geographic mechanisms of diversification (Bennett & O’Grady, 2013). Our network analysis indicates that specialization and modularity are more pronounced on Maui than on Hawaii Island (Figs. 3, 4). This observation is consistent with the suggestion from the *Nesophrosyne* phylogenetic results and suggests that an approximate age of 1 million years may be necessary for host plant specialization to become the dominant process in the sequence of diversification. Available dating analyses of other arthropod taxa indicate that genetic structure can develop in much less than 1 million years (Table 1), and suggests that landscape processes may dominate the earliest stages of diversification across taxa in the Hawaiian Islands. Other taxa at low trophic levels, such as the herbivorous *Trioza* psyllids, detritivorous *Laupala* crickets and fungivorous *Drosophila*, show similar signals of geographic isolation combined with ecological and sexual processes driving genetic divergence and diversification across sites as young as those on the Hawaii Island (Percy, 2003; Percy *et al*. 2006; Mendelson & Shaw, 2005; Magnacca *et al.*, 2008; O’Grady *et al.*, 2011). By contrast, spiders, which are predatory, develop genetic discontinuities at larger spatial and temporal scales. (Table 1). Further work is needed to determine if this pattern of less genetic differentiation in predators compared to herbivores is general.

## Macroecological metrics: Network structure and steady state

Across the Hawaiian archipelago, nestedness appears to decrease generally with site age, and is highest on the geologically youngest volcano, Kilauea. The high nestedness on Kilauea, is likely a result of new species arriving by immigration that have a high probability to eat or be eaten by the generalist species already present at the site (Bascompte & Jordano, 2007). However, despite the high nestedness on Kilauea, we do not find Kilauea to conform to the statistical steady state predication of METE. The observed deviations from METE at Kilauea appear to be largely driven by a surplus of singleton links (Fig. 3), which may reflect a state of “incomplete” assembly, possibly by lower species richness of the plant and herbivore biotas. Conversely, at Kohala, at intermediate age (150 ky), observations are not significantly different from predictions from METE. We contend that the model fits well because the site has had sufficient time to undergo ecological succession and thus arrive at a statistical steady state but is still too young to be affected by ecological specialization and rapid *in situ* diversification associated with host plants.

Interestingly, the observations of communities on older Maui and Kauai sites show strong deviations from expectations of METE (Fig. 4). METE is agnostic about what mechanisms are determining the values of the state variables that lead to its macroecological predictions (Harte, 2011). Indeed, the use of maximum entropy in ecology is inspired by its application to physical systems, in which change through time is simple. It does not account for the evolutionary history of biological systems, which are potentially very different due to the complex change through time produced by speciation, extinction and adaption to novel ecosystems. Thus, deviations from theory need to be interpreted in the context of the study design. We hypothesize that the influence of evolutionary assembly on these biotas is what drives them away from and ecological statistical steady state. One possible explanation for the strong deviations from expectations of METE, compared with observations at our intermediate-aged site (Kohala), is that while the ages of Maui and Kauai are sufficient for evolutionary assembly driven by specialization and diversification on host plants, the older island age of these islands may have led to range contractions and possibly extinction of plant species on the oldest island of Kauai (Whittaker et al 2008). METE is currently being extended to incorporate speciation and migration, which will allow testing of these ideas more directly in the future.

Our results show decreased nestedness and increased modularity on Maui and Kauai. Coevolution between interacting species should lead to greater modularity (Donatti *et al.*, 2011; Nuismer *et al.*, 2013). However, the influence of certain network properties such as nestedness on stability is still unknown, and so theoretical predictions of how network properties should change over evolutionary time, generally, are lacking. Theoretical and empirical studies have suggested that nestedness may or may not promote stability (Allesina & Tang 2012; Suweis *et al*. 2014). Furthermore, almost all studies of food webs have focused primarily on single, or short ecological timespans of network development that do not span as large an amount of evolutionary time as included here (e.g., Albrecht et al 2010). Food webs are dynamic emergent entities, with broad topological characteristics that may change dramatically over time (e.g., Yeakel et al 2013). To our knowledge, our study represents the first to evaluate network topology over larger temporal scales, and we argue that age-structured landscapes such as the Hawaiian archipelago are promising for resolving longstanding debates on the causes and consequences of network properties such as nestedness.

We found that single-island endemics were always more specialized than multiple-island endemics. While dietary breadth has been positively associated with geographic range size, the direction of causality is unclear (Slatyer et al 2013). Dietary breadth may allow some species to colonize other islands, but dietary breadth may also be driven by local adaptation to exploit locally abundant hosts across a large range. Nevertheless, both scenarios are consistent with the hypothesis that *in situ* formation of single-island endemics may be the product of coevolution and specialization. At the Kohala site, which showed the best fit to maximum entropy theory, single-island endemic and multiple-island endemic species alike show increased generalization (i.e. higher degree, or more links; Fig 3), while at the youngest site Kilauea, specialist single-island endemics may be limited by low plant diversity and thus appear more specialized (Fig 3). Conversely at the oldest site on Kauai, where plant diversity is high (Kitayama & Mueller-Dombois, 1995), single-island endemics again are associated with decreased degree and thus genuine specialization (Fig. 3). On Maui, single-island endemics show statistically significant increases in generalization but this pattern disappears when analyzing the data at the resolution of plant genera, thus suggesting that Hemiptera species endemic to Maui may benefit from the diversification of plant species within genera.

## Future Research

The data and analyses presented here describing insect and plant communities across a chronosequence of habitats in Hawaii generate testable hypotheses concerning the relative importance of ecological and evolutionary processes in community assembly. Our work to date suggests the overarching hypothesis that ecological processes dominate community assembly in younger environments, with evolutionary processes becoming more important as communities age. We can also make predictions about the sequence of community assembly based on proposed mechanisms.

In younger communities we predict characteristics of ecological assembly, such as those strongly influenced by immigration. Species in such communities should resemble random samples from regional source pools and thus metrics describing these communities will approach expectations of ecological statistical steady. An exception will be communities still undergoing initial stages of primary succession, which will change rapidly through time and represent non–random samples of source pools. We also predict that these communities will exhibit a nested network structure, assuming new species will eat or be eaten by the generalist species already present in the community, as suggested by previous work on nestedness (Bascompte & Jordano, 2007) and by our finding that widespread species tend to be generalists (Fig. 4).

Following the same logic, in older communities we expect to see characteristics of evolutionary assembly, dominated by processes such as niche exploration, adaptation and speciation. In such communities we predict a higher level of specialization and modularity (Bascompte & Jordano, 2007; Donatti *et al.*, 2011; Nuismer *et al.*, 2013).

### Ecological data: Assembly of species into communities

In order to build a more rigorous understanding of the assembly process in both younger and older communities, fine-grained sampling of all macroscopic arthropod taxa is needed from a large number of sites across the island chronosequence. This will allow an assessment of changes in overall species composition and diversity across all players in the time-calibrated landscape (*sensu* Gruner, 2007). Such data will allow us to test how entire arthropod communities of different aged substrates deviate from statistical steady state as predicted by METE (Harte, 2011). For example, predators, whose assemblages are likely more dominated by immigration and ecological assembly (Fig. 2) may never show strong deviations from METE predictions whereas herbivores could show increasing deviation with age in agreement with the network results of this paper (Fig. 3).

### Evolutionary data: Diversification within species

The current study demonstrates that taxa from different trophic guilds differ in the scale at which differentiation occurs and highlights the importance of fragmentation of the landscape in facilitating differentiation. Future work will be aimed at gathering data for additional focal taxa within this system spanning different trophic levels. We will use these data to understand taxonomic differences in the rate of differentiation, to assess the roles of genetic fusion and fission and the spatial scale over which they are important (Gillespie and Roderick 2014), and to detail the relative rates of speciation and extinction across the island chronosequence.

## CONCLUSIONS

NEED ONE PARAGRAPH HERE SAYING HOW WE SHOWED THIS (from intro): Here we show how the effects on diversity of changing interactions across an ecological-evolutionary continuum can be identified by studying the biota of archipelagoes with islands sequentially arranged by age combined with predictions from idealized ecological theories (such as the unified neutral theory [Hubbell, 2001] or the maximum entropy theory of ecology [Harte 2011]). The different aged islands provide approximate snapshots of the eco-evolutionary process at different stages. Ecological theory provides an idealized “null” baseline against which to compare observed patterns.

We found that taxa in the lower trophic levels developed genetic structure even in the youngest habitats of the observed chronosequence and at smaller spatial scales (Table 1, Fig. 2). Thus, lower trophic levels are affected by *in situ* processes of diversification very early in the chronosequence, compared to higher trophic levels. Network nestedness decreased while modularity increased with age (Fig. 4), again indicating a possible shift from assembly driven by *ex situ* immigration early on, to one based on *in situ* diversification, such as in co-diversification of insect herbivores with host plants (Bascompte & Jordano, 2007; Donatti *et al.*, 2011). That single island endemics (likely the product of *in situ* diversification) show more specialization at older sites compared to more broadly distributed species (those taxa more likely to be initial colonists; Fig. 3) also supports this hypothesis.

This study provides a framework for using chronologically arranged oceanic island systems to examine the interplay between evolutionary and ecological processes in shaping biodiversity. Our initial results provide a clear hypothesis that ecological processes dominate community assembly in younger environments, with evolutionary processes becoming more important as communities age. We demonstrate how this approach can provide insights into the development of communities over ecological–evolutionary time, and the dynamic feedbacks involved in assembly.

# Acknowledgements

We are indebted to many scientists and land managers in Hawaii that have provided access to the lands: Pat Bily (The Nature Conservancy of Hawaii), Melissa Dean, Christian Giardina, and Tabetha Block (Hawaii Experimental Tropical Forests), Betsy Gagne (Natural Area Reserve System), Lisa Hadway and Joey Mello (Department of Forestry and Wildlife Hilo), Cynthia King and Charmian Dang (Department of Land and Natural Resources), and Rhonda Loh (Hawaii Volcanoes National Park). We thank Robert Ricklefs, Lauren Ponisio and Anna Hiller for thoughtful commentary. We are very grateful to Guida Santos, Richard Field and Robert Ricklefs for inviting us to contribute to this special issue. The research was supported by the National Science Foundation DEB 1241253.

# Biosketch

The authors are part of a large team involved in understanding how biodiversity and complex ecosystems emerge from ecological and evolutionary processes. The project, funded by the National Science Foundation’s “Dimensions in Biodiversity”, focuses on the geological chronosequence provided by the Hawaiian islands.

# 

# TABLES

**Table 1.** Results of the analyses of molecular variance (AMOVA) that partitions molecular genetic variation between volcanoes and between sites within volcanoes for arthropod lineages found within the study sites on Hawaii Island. Where estimates of divergence through molecular dating are available for the taxa, they are presented to show the timeframe within which this genetic structure has developed.

# Figure captions

**Figure 1.** Substrate age map of the islands of Kauai, Maui and Hawaii Island. Colors correspond to substrate age from young (yellow) to old (blue). Focal sites are shown as black circles (on Hawaii Island, Kohala is in the north, Kiluea in the south) while sampling sites for genetic data are represented by gray circles.

**Figure 2.** Genetic population structure () among sites within volcanoes with volcano age for insects and spiders. Calculations were based on mitochondrial DNA only (see Table 1 for details). The plant-feeding groups, specifically the sap-feeding Hemiptera, show higher genetic structure among sites on young volcanoes relative to older volcanoes whereas detritivores (crickets), fungivores (*Drosophila*), and in particular predators (spiders), show little structure on young volcanoes. For spiders, substantial structure develops only later in the chronosequence, e.g., on Maui at approximately 1 million years. Numbers refer to different species: 1. *Nesosydne chambersi*, 2. *Nesosydne raillardiae*, 3. *Nesosydne bridwelli*, 4. *Trioza HB*, 5. *Trioza HC*, 6. *Drosophila sproati*, 7. *Laupala cerasina*, 8. *Tetragnatha anuenue*, 9. *Tetragnatha brevignatha*, 10. *Tetragnatha quasimodo*, 11. *Theridion grallator*.

**Figure 3.** Patterns in degree distributions across sites, comparing archipelago-wide endemics (Cosmopolitans) with single island endemic (Endemics) taxa. Top panels show that networks deviate most from METE predictions on youngest and oldest sites. Inset figures show the distribution of simulated mean squared errors; if the red line falls within the gray region (95% confidence interval) the data are not significantly different from the predictions of maximum entropy theory; all sites except Kohala deviate from the predications. The bottom panel shows the number of links for endemics versus cosmopolitans. Endemics show lower linkage overall, but significantly increase on the intermediate-aged site Maui (highlighted with dotted box). Kohala shows increased linkage overall (highlighted with solid box).

**Figure 4.** Trends in network metric nestedness and modularity through time. Nestedness decreases while modularity increases. Error bars represent 95% confidence intervals from a null model simulation.

# References

Albrecht, M., Riesen, M. & Schmid, B. (2010) Plant-pollinator network assembly along the chronosequence of a glacier foreland. *Oikos*, **119**, 1610–1624.

Allesina, S. & Tang, S. (2012) Stability criteria for complex ecosystems. *Nature*, **483**, 205–208.

Almeida-Neto, M., Gumarães, P., Gumarães, P., Loyola, R. & Ulrich, W. (2008) A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos*, **117**, 1227–1239.

Bascompte, J. & Jordano, P. (2007) Plant-animal mutualistic networks: the architecture of biodiversity. *Annual Review of Ecology and Systematics*, **38**, 567–93.

Bennett, G.M. & O’Grady, P.M. (2013) Historical biogeography and ecological opportunity in the adaptive radiation of native Hawaiian leafhoppers (Cicadellidae: *Nesophrosyne*). *Journal of Biogeography*, **40**, 1512–1523.

Borer, E.T., Seabloom, E.W., Gruner, D.S., Harpole, W.S., Hillebrand, H., Lind, E.M., Adler, P.B., Alberti, J., Anderson, T.M., Bakker, J.D., Biederman, L., Blumenthal, D., Brown, C.S., Brudvig, L.A., Buckley, Y.M., Cadotte, M., Chu, C., Clelamd, E., Crawley, M.J., Daleo, P., Damschen, E.I., Davies, K.F., DeCrappeo, N.M., Du, G., Firn, J., Hautier, Y., Heckman, R.W., Hector, A., HilleRisLambers, J., Iribarne, O., Klein, J.A., Knops, J.M.H., La Pierre, K.J., Leakey, A.D.B., Li, W., MacDougall, A.S., McCulley, R.L., Melbourne, B.A., Mitchell, C.E., Moore, J.L., Mortensen, B., O’Halloran, L.R., Orrock, J.L., Pascual, J., Prober, S>M., Pyke, D.A., Risch, A.C., Schuetz, M., Smith, M.D., Stevens, C.J., Sullivan, L.L., Williams, R.J., Wragg, P.D., Wright, J.P. & Yang, L.H.(2014) Herbivores and nutrients control grassland plant diversity via light limitation. *Nature*, **508**,517–520.

Brewer, M.S., Carter, R., Croucher, P.J.P & Gillespie, R.G. (2015) Shifting Habitats, Morphology and Selective Pressures: Developmental Polyphenism in an Adaptive Radiation of Hawaiian Spiders. *Evolution*, **69**, 162-178

Brose, U., Williams, R.J. & Martinez, N.D. (2006) Allometric scaling enhances stability in complex food webs. *Ecology Letters*, **9**, 1228–1236.

Cavender-Bares, J., Kozak, K.H., Fine, P.V. & Kembel, S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters*, **12**, 693–715.

Csardi, G. & Nepusz, T. (2006) The igraph software package for complex network research. *InterJournal*, *Complex Systems*, 1695.

Donatti, C.I., Guimarães, P.R., Galetti, M., Pizo, M.A., Marquitti, F. & Dirzo, R. (2011) Analysis of a hyper-diverse seed dispersal network: modularity and underlying mechanisms. *Ecology Letters*, **14**, 773–781.

Eldon, J., Price, J.P., Magnacca, K. & Price, D.K. (2013) Patterns and processes in complex landscapes: testing alternative biogeographical hypotheses through integrated analysis of phylogeography and community ecology in Hawai’i. *Molecular Ecology*, **22**, 3613–3628.

Excoffier, L. & Lischer, H.E. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under linux and windows. *Molecular Ecology Resources*, **10**, 564–567.

Dewar, R.C., and Porté, A. (2008) Statistical mechanics unifies different ecological patterns. *Journal of Theoretical Biology.* 251, 389-403.

Gillespie, R. (2004) Community assembly through adaptive radiation in hawaiian spiders. *Science*, **303**, 356–359.

Gillespie, R.G. & Roderick, G.K. (2014) Geology and climate drive diversification. *Nature*, **509**, 207–298.

Gillespie, R.G. & Baldwin, B.G. (2009) Island biogeography of remote archipelagoes. In *The Theory of Island Biogeography Revisited* (Losos, J.B. & Ricklefs, R.E.; eds.), Princeton University Press, Princeton, NJ. pp 358–387.

Goodman, K.R., Kelley, J.P., Welter, S.C., Roderick, G.K. & Elias, D.O. (2015) Rapid diversification of sexual signals in Hawaiian Nesosydne planthoppers (Hemiptera: Delphacidae): the relative role of neutral and selective forces. *Journal of Evolutionary Biology***. 28**, 415–427

Goodman, K., Welter, S. & Roderick, G. (2012) Genetic divergence is decoupled from ecological diversification in the Hawaiian *Nesosydne* planthoppers. *Evolution*, **66**, 2798–2814.

Gruner, D.S. (2007) Geological age, ecosystem development, and local resource constraints on arthropod community structure in the Hawaiian Islands. *Biological Journal of the Linnean Society*, **90**, 551–570.

Gull, S. F., and Newton, T. J. (1986). Maximum entropy tomography. *Applied Optics,* 25, 156–160.

Harte, J. (2011) *Maximum Entropy and Ecology*, Oxford University Press, Oxford, England.

Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton University Press, Princeton, NJ.

Jaynes, E.T. (1957) Information theory and statistical mechanics. *Physical Reviews*, **106**, 602.

Jetz, W., Thomas, G., Joy, J., Hartmann, K. & Mooers, A. (2012) The global diversity of birds in space and time. *Nature*, **491**, 444–448.

Golan, A., Judge, G., and Miller, D. (1996). *Maximum entropy econometrics: Robust estimation with limited data*. John Wiley and Sons, New York

Kitayama, K. & Mueller-Dombois, D. (1995) Vegetation changes along gradients of long-term soil development in the Hawaiian montane rainforest zone. *Vegetatio*, **120**, 1–20.

Magnacca, K.N., Foote, D. & O’Grady, P.M. (2008) A review of the endemic Hawaiian Drosophilidae and their host plants. *Zootaxa*, **1728**, 1–58.

Mendelson, T.C. & Shaw, K.L. (2005) Sexual behaviour: rapid speciation in an arthropod. *Nature*, **433**, 375–376.

Miller, M. (1997) Tools for population genetic analysis (tFPGA), 1.3: a windows program for the analysis of allozyme and molecular population genetic data. distributed by the author: http://www.marksgeneticsoftware.net/tfpga.htm.

Newman, M. & Girvan, M. (2004) Finding and evaluating community structure in networks. *Physical Review E*, **69**, 026113.

Nuismer, S.L., Jordano, P. & Bascompte, J. (2013) Coevolution and the architecture of mutualistic networks. *Evolution*, **67**, 338–354.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O’Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H. & Wagner, H. (2013) *vegan: Community ecology package*.

Olesen, J., Bascompte, J., Dupont, Y. & Jordano, P. (2007) The modularity of pollination networks. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 19891–19896.

O’Grady, P.M., Lapoint, R.T., Bonacum, J., Lasola, J., Owen, E., Wu, Y. & DeSalle, R. (2011) Phylogenetic and ecological relationships of the Hawaiian *Drosophila* inferred by mitochondrial DNA analysis. *Molecular Phylogenetics and Evolution*, **58**, 244–256.

Percy, D.M. (2003) Radiation, diversity and host plant interactions among island and continental legume-feeding psyllids. *Evolution*, **57**, 2540–2556.

Percy, D.M., Taylor, G.S. & Kennedy, M. (2006) Psyllid communication: acoustic diversity, mate recognition and phylogenetic signal. *Invertebrate Systematics*, **20**, 431–445.

Phillips, S. J., *et al.* (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling.*, 190, 231-259.

Price, J. & Clague, D. (2002) How old is the Hawaiian biota? Geology and phylogeny suggest recent divergence. *Proceedings of the Royal Society London B Biological Sciences*, **269**, 2429–2435.

Price, J.P. (2012) *Mapping plant species ranges in the Hawaiian Islands: developing a methodology and associated GIS layers*, US Department of the Interior, US Geological Survey.

R Core Team (2014) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ricklefs, R.E. (2004) A comprehensive framework for global patterns in biodiversity. *Ecology Letters*, **7**, 1–15.

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

Roderick, G.K. & Percy, D.M. (2008) Host plant use, diversification, and coevolution: Insights from remote oceanic islands. In *Specialization, Speciation, and Radiation: The Evolutionary Biology of Herbivorous Insects* (K. Tilmon, ed.), pp. 151–161. University of California Press, Berkeley.

Rosindell, J. & Phillimore, A.B. (2011) A unified model of island biogeography sheds light on the zone of radiation. *Ecology Letters*, **14**, 552–560.

Roussev, V. (2010). Data fingerprinting with similarity digests. In *Advances in Digital Forensics VI* (Chow, K.-P. & Shenoi, S., eds.), pp. 207-226. Springer Berlin Heidelberg.

Sherrod, D.R., Murai, T. & Tagami, T. (2007) New K–Ar ages for calculating end-of-shield extrusion rates at west Maui volcano, Hawaiian Island chain. *Bulletin of Volcanology*, **69**, 627–642.

Slatyer, R.A., Hirst, M. & Sexton, J.P. (2013) Niche breadth predicts geographical range size: a general ecological pattern. *Ecology Letters*, **16**, 1104–1114.

Suweis, S., Simini, F., Banavar, J.R. & Maritan, A. (2014) Emergence of structural and dynamical properties of ecological mutualistic networks. *Nature,* **500**, 449–452.

Thompson, J. N., C. Schwind, P. R. Guimarães, Jr. & M. Friberg. (2013) Divergence through multitrait evolution in coevolving interactions. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 11487–11492.

Tilman, D. (2004) Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 10854–10861.

Ulrich, W., Almeida-Neto, M. & Gotelli, N.J. (2009) A consumer’s guide to nestedness analysis. *Oikos*, **118**, 3–17.

Vandergast, A., Gillespie, R. & Roderick, G. (2004) Influence of volcanic activity on the population genetic structure of Hawaiian *Tetragnatha* spiders: fragmentation, rapid population growth, and the potential for accelerated evolution. *Molecular Ecology*, **13**, 1729–1743.

Vázquez, D., Poulin, R., Krasnov, B. & Shenbrot, G. (2005) Species abundance and the distribution of specialization in host-parasite interaction networks. *Journal of Animal Ecology*, **74**, 946–955.

Vázquez, D.P., Blüthgen, N., Cagnolo, L. & Chacoff, N.P. (2009) Uniting pattern and process in plant-animal mutualistic networks: a review. *Annals of Botany*, **103**, 1445–1457.

Vitousek, P.M. (2004) *Nutrient Cycling and Limitation: Hawai‘i as a Model System.* Princeton University Press, Princeton, NJ.

Wagner, W. & Funk, V. (1995) *Hawaiian Biogeography: Evolution on a Hot Spot Archipelago*, Smithsonian Institution Press, Washington, DC.

Warren, B.H., Simberloff, D., Ricklefs, R.E., Aguilée, R., Condamine, F.L., Gravel, D., , Morlon, H., , Mouquet, N., Rosindell, J., Casquet, J., Conti, E., Cornuault, J., Fernández-Palacios, J.M. Hengl, T., Norder, S.J., Rijsdijk, K.F., Sanmartín, I., Strasberg, D., Triantis, K., Valente, L.M., Whittaker, R.J., Emerson, B.C., Gillespie, R.G. & Thébaud, C. (2015) Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecology Letters*, **18**: 200-217.

Whittaker, R.J., Triantis, K.A. & Ladle, R.J. (2008) A general dynamic theory of oceanic island biogeography. *Journal of Biogeography*, **35**, 977–994.

Wiens, J.J. (2011) The causes of species richness patterns across space, time, and clades and the role of “ecological limits”. *The Quarterly Review of Biology*, **86**, 75–96.

Williams, R.J. (2010) Simple MaxEnt models explain food web degree distributions. *Theoretical Ecology*, **3**, 45–52.

Williams, R.J. & Martinez, N.D. (2000) Simple rules yield complex food webs. *Nature*, **404**, 180–183.

Yeakel, J.D., Guimarães Jr, P.R., Bocherens, H. & Koch, P.L. (2013) The impact of climate change on the structure of Pleistocene food webs across the mammoth steppe. *Proceedings of the Royal Society B Biological Sciences*, **280**, 20130239.

Additional references to the data sources may be found in Appendix S1 at [URL].