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, G. M Bennett5, M. S. Brewer1, D. D. Cotoras2, C. P. Ewing4, J. Harte1, N.D. Martinez3, P. M. O’Grady1, D. M. Percy6, D. K. Price4, G. K. Roderick1, K. L. Shaw7, D. S. Gruner8#, R. G. Gillespie1#

1. Environmental Science, Policy, and Management, University of California, Berkeley, California 94720-3114
2. Integrative Biology, University of California, Berkeley, California 94720-3140
3. Pacific Ecoinformatics and Computational Ecology Lab, Berkeley, California 94703
4. Biology, University of Hawaii, Hilo, Hawaii, 96720-4091
5. Integrative Biology, University of Texas, Austin, Texas 78712
6. Entomology, The Natural History Museum, London, UK SW7 5BD
7. Neurobiology and Behavior, Cornell, Ithaca, New York 14853-7601
8. Department of Entomology, University of Maryland, College Park, Maryland 20742-4454
9. \*Contributed equally; # co-senior authors;
10. 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: 299

Number of words in main body of the paper: 4973

Number of references: 50

#### Aim

Understanding how ecological and evolutionary processes synergistically determine biodiversity patterns remains a central goal in biology. Using ecological theory as a conceptual guide and data from multiple arthropod lineages, we use the isolated archipelago of Hawaii to explore how complex communities emerge from the interplay of ecological (population dynamics, dispersal, trophic interactions) and evolutionary (genetic structuring, adaptation, speciation, extinction) processes.

#### Location

The Hawaiian Islands (19.5°N, 155.5°W).

#### Methods

To infer processes involved in early diversification we synthesize data on genetic 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 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 time perspective provided by the island chronosequence and genetic information, we demonstrate that species in lower trophic levels develop population genetic structure more quickly than species of higher trophic levels. Higher trophic levels also show endemism, though it evolves more slowly and over larger areas. Moreover, in analyzing plant-herbivore networks across an increasing substrate age gradient we find trends of higher specialization and increasing deviation from the statistical steady state expected from theoretical predictions of food web structure.

#### Main conclusions

We show how ecological theory can leverage natural experiments on oceanic islands of known chronologies to understand the interplay between evolutionary and ecological processes in shaping present-day biodiversity. We advocate for combining perspectives gained from coupled molecular and community-level data analyzed in the context of ecological theory. We further show the utility of applying theory in a chronosequence context to better illuminate the interplay of ecological mechanisms, speciation, extinction and adaptation in driving contemporary biodiversity patterns.

Introduction

Contemporary biodiversity is a product of speciation, extinction and dispersal all conditioned by ecological interactions with the biotic and abiotic environment. Because these processes occur on different temporal and spatial scales, disentangling the relative influence of local ecological mechanisms from evolutionary and historical processes is challenging (Ricklefs, 2004). Archipelagoes with islands of different size and age enable us to sort out the effects on diversity of processes operating on different temporal and spatial scales.

The integration of ecological theory with an evolutionary perspective has the potential to reveal dynamics that generate biodiversity. The evolutionary processes of speciation and extinction tend to be viewed as regulating regional species pools, occurring in a manner largely removed from local ecology (Hubbell, 2001; Cavender-Bares *et al.*, 2009; Wiens, 2011). Ecological mechanisms 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). Recent theoretical advances have further refined the causes and consequences of ecological drift (Hubbell, 2001; Rosindell & Phillimore, 2011), re-vitalized classical niche-based mechanisms such as niche partitioning (Tilman, 2004), competition and predation (Borer *et al.*, 2014), and put species interactions in a network theoretic context (Williams & Martinez, 2000; Brose *et al.*, 2006).

New 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). Here, we propose an integrative framework that harnesses both sources of advancement placed in the context of age-structured archipelagos to study evolutionary community assembly. We argue that mechanistically simplified ecological theories (such as neutral theory CITE or the maximum entropy theory of ecology CITE) can be used as powerful null models to identify ecological assemblages with unique underlying biological processes warranting further study. Age-structured island systems combined with population genetic and phylogenetic data can reveal whether it is the evolutionary history behind such communities that drives their deviation from theoretical expectations. We provide an initial case study using a synthesis of published data on arthropod lineages in the Hawaiian Islands. We estimate metrics of evolutionary and ecological dynamics across communities found in geologic settings that range in age from 500 yr to 5 myr. We estimate taxon-specific timelines for the development of genetic structure. This is contextualized with macroecological metrics of community structure, using predictions from statistical equilibrium and ecological network theory to examine changes over the island chronosequence.

## Hotspot oceanic archipelagos as model systems

Hotspot oceanic islands are opportune model systems for studying the interplay of local ecological mechanisms and large–scale, historical, 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 . We hypothesize that the contributions of evolutionary and ecological assembly will vary according to geological age of the environment, taken as an indicator of the total time communities have had to assemble and over which *in situ* diversification could occur. Age–structured, hot–spot island archipelagoes thus have the potential to stratify the eco–evolutionary process of community assembly.

Younger communities originate mostly from initial immigration and thus should be dominated by classical ecological mechanisms operating on a source pool whose evolution is removed from the local setting. Conversely, older islands could allow for the combined interaction and feedback of diversification and local ecological dynamics. Moreover, because dispersal, and hence connectivity between sites, differs between taxa, the relative role of evolutionary and ecological assembly will differ between taxa. While many archipelagos around the world share these biotic and geologic properties, the Hawaiian archipelago provides a particularly useful system for study because its geological chronology (Price & Clague, 2002) and patterns of biodiversity are well characterized (Wagner & Funk, 1995).

## Development of genetic structure

If dispersal and gene flow are high then local adaptation and *in situ* diversification are less important in shaping community composition. By contrast, when connectivity is low genetic divergence will be high and the ecologies of distinct populations in different localities are free to diverge through local selective pressures and drift leading to the possibility of diversification. For these reasons, the magnitude of genetic connectivity among populations provides a measure of the relative importance of dispersal-driven assembly and assembly by in situ diversification in determining community composition. Here, by using the chronosequence, we can analyze populations from multiple sets of taxa across trophic guilds occurring in geological contexts from young to old. If the evolutionary histories and diversification processes experienced by these taxa are important in determining unique biological assemblages we would expect the magnitude of deviations from a null model—provided by simplified ecological theory—would be correlated with age along chronosequence and/or degree of population structure.

## Macroecological metrics

While we expect the mechanisms underlying the generation and maintenance of biodiversity to change across chronological sequences, studies to date have rarely moved beyond reporting basic patterns (Gillespie & Baldwin, 2010). Theories of biodiversity (e.g. Hubbell, 2001; Harte, 2011) provide a necessarily simplified view of biodiversity and deviations from theory can provide evidence of which particular ecological patterns require more biologically realistic mechanisms for their explanation and which patterns are consistent with statistical chance.

The Maximum Entropy Theory of Ecology (METE; Harte, 2011) provides a prediction for idealized ecological communities in statistical steady state, the situation in which a system’s behavior is governed by a simple set of state variables and no further mechanisms are required. 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 [1,2], economics [3], forensics [4], imaging technologies [5-7], and recently ecology [8-15]. 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). Least biased in this case means making no further assumptions about the system other than those necessary to meet the constraints of the state variables. This results in predicted distributions that are as uniform as possible, and thus contain the maximum possible information entropy. Put another way these predictions are the most ignorant possible predictions about the system.

Studying the unique ecological conditions and evolutionary histories of real world deviations from METE can provide insights into the processes driving ecological systems away from this statistical steady state and toward alternate system states (Harte, 2011). We expect that different aged communities along the Hawaiian chronosequence will deviate differently from METE, because we hypothesize the processes of speciation, extinction, adaptation and colonization may themselves drive Hawaiian communities out of statistical steady state.

METE can successfully predict various metrics of an ecological community (Harte, 2011), including network metrics that describe trophic interactions between species (Williams, 2010; Harte, 2011). Ecological network theory derives from evolutionary concepts such as coevolution (Bascompte & Jordano, 2007; Donatti *et al.*, 2011; Nuismer *et al.*, 2013; Thompson et al. 2013) and so presents an opportune starting place to study ecological and evolutionary feedbacks. The distribution of linkages in ecological networks has been used to determine whether plant-animal interaction networks assemble neutrally or through deterministic processes (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). This means that species with fewer interactions (i.e. more specialized species) will interact with a subset of the species that generalists interact with. In this way interaction nestedness is mathematically equivilant to island nestedness (in which less species rich islands are subsets of more species rich islands; CITE). However, we only consider network nestedness here.

In this paper we integrate methods from population genetics to theoretical ecology using the chronosequence of the Hawaiian Archipelago to understand the nexus between ecological and evolutionary forces community assembly. Moving from young to old across the chronosequence we evaluate (1) the rate and pattern of genetic connectivity among populations of taxa from different trophic levels as they diversify from populations to form new species; (2) the processes underlying the structure of species interaction networks given the backdrop of population divergence; and (3) the processes involved in diversification as species form and accumulate and how this dynamic drives deviations from statistical steady state. We use data (mostly published) on population genetic structure and species interactions as a proof of concept. With this framework, our goal is to show how communities develop over ecological-evolutionary time, and the dynamic feedbacks involved assembly.

# Methods

## Hawaii as an eco-evolutionary study system

The geological landscape of the Hawaiian Islands offers a matrix of volcanic substrates mapped in fine detail by chronological age and geochemical composition (Sherrod *et al.*, 2007). Ongoing volcanic activity has created a dynamic mosaic of habitats with transitory to long-lasting fragmentation within landscapes. Isolation on scales of hundreds to thousands of meters and hundreds to thousands of years can be sufficient for genetic differentiation among some arthropod populations among habitats (Goodman *et al.*, 2012; Bennett & O’Grady, 2013; Eldon *et al.*, 2013), while insufficient to isolate others (Vandergast *et al.*, 2004). On larger spatial and temporal scales, distinct volcanoes and islands, with their semi-independent histories of evolutionary community assembly from recent to ancient, comprise a geological chronosequence spanning from present day up to 5 million years across Hawaii Island to Kauai. This chronosequence offers a cross-sectional study system in which younger islands provide a proxy for the early evolutionary development that older islands likely underwent as well. Key to this opportunity is the fact that many endemic taxa immigrated to and initially diversified on ancient Hawaiian islands (now no more than atolls; CITE) and so have already experienced the unique evolutionary pressures of survival on islands (CITE).

## To investigate how ecological patterns change in response to varied evolutionary contexts we selected four focal sites across the chronosequence of substrate 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 climate (mean annual precipitation 2000-3000 mm).

## Compilation and analysis of genetic data

To evaluate the balance between regional immigration and potential for local ecological differentiation, we measured how molecular variation is partitioned within species within 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:

1. Herbivorous *Nesosydne* planthoppers (COI and microsatellites; Goodman *et al.*, 2012 GenBank accession numbers XXX-XXX); *Trioza* psyllids (COI, cytB; GenBank accession numbers XXX-XXX); and fungivorous *Drosophila sproati*: (COII; Eldon *et al.*, 2013) that maintains tight host plant associations.
2. Detritivorous *Laupala* crickets (AFLPs; Mendelson *et al.*, 2004); and
3. several predatory spiders species (COI and allozymes; Roderick *et al.*, 2012)

In the case of previously unpublished molecular data, sequences for *Nesosydne* planthoppers were generated following protocols described in (Goodman *et al.*, 2012) and sequences for *Trioza* psyllids were generated following protocols described in Percy (2003). Because these are existing data, the samples used in this analysis were not necessarily collected within the focal sites described above for the ecological analysis. They were collected from across Hawaii Island and Maui and are used here to provide an estimate of how arthropod populations have accumulated genetic diversity and divergence within the dynamic landscape of the focal sites.

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 Maui Nui. 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 evaluate the role of landscape age in allowing *in situ* genetic diversity and potential for 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 of plant-herbivore networks

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 hemipteran herbivores endemic to the archipelago 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 Hemiptera 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 hemipteran species 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 was 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 hemipteran species presences and the other using the less conservative data set.

## Analysis of plant-herbivore networks

We hypothesize that communities differentially depart from statistical steady state along the continuum form those dominated by ecological processes to those with potential complex evolutionary feedbacks. We used METE (Williams, 2010; Harte, 2011) to compute the statistical steady state for the hemipteran degree distribution (distribution of the number of plant hosts to each hemipteran species). 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 used for METE estimation and Monte Carlo methods are available in the supplement.

To further investigate how *in situ* diversification leaves a potentially unique signature on network structure we analyzed the number of links assigned to each hemipteran species (the degree distribution) separately for 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). 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, which describes the degree of asymmetry of species interactions connecting specialists and generalists (Bascompte & Jordano, 2007; Ulrich *et al.*, 2009), and modularity which describes the degree to which interactions are concentrated within subsets of species but not between subsets (Newman & Girvan, 2004; Olesen *et al.*, 2007).

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 discontinuity among populations

The analysis of molecular variance (AMOVA) revealed evidence of significant population genetic structure from the smallest to the largest spatial scales examined, all within a very recent timeframe. For mitochondrial loci, the amount of statistically significant molecular variation partitioned to among sites within volcanoes ranged from 0.037–0.92 and to the 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 (larger observed variance within volcanoes may represent the effects of colonization from multiple, isolated areas or sampling error). 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 on young volcanoes relative to among localities on older volcanoes (Fig. 2). By contrast, predatory spiders exhibited higher genetic population structure only on older volcanoes (e.g. Maui).

The observed levels of genetic divergence have evolved rapidly. For example, within species genetic divergence in planthoppers has evolved in as little as 2,600 years (Goodman *et al.*, 2012). 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 developed subsequently (Table 1). That the estimates are older than the known age of the Big Island, suggest that divergence in DNA now reflected in different lineages pre-dated their colonization of the Big Island, or alternatively may be due to sampling error.

## Evolving network structure

The Hemiptera species degree distribution 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 middle aged site of Kohala, minor deviations from maximum entropy are no different than expected by chance indicating the Kohala Hemiptera assemblage matches the predictions of maximum entropy.

The generalized linear model revealed that there are also 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. Further, single-island endemics become significantly more generalist on the middle aged Maui site; however this pattern disappears when analyzing links to plant genera instead of species. The slightly younger Kohala shows increased generalization overall. When considering the degree distribution defined by trophic links to plant genera instead of plant species, the pattern of increased generalization holds for the 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 intra-genus plant diversity on that island.

Network nestedness decreased with age while modularity increased (Fig. 4). This trend is found 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

Taxa in the lower trophic levels, as compared to higher trophic guilds, developed genetic discontinuities more quickly along the chronosequence and at much smaller spatial scales (Table 1, Fig. 2), allowing them the opportunity to diverge ecologically. Network nestedness decreased while modularity increased with age (Fig. 4), indicating a possible shift from assembly driven by *ex situ* immigration early on, to one based on *in situ* co-diversification with host plants (Bascompte & Jordano, 2007; Donatti *et al.*, 2011). This possibility is further strengthened by the observation that single island endemics (likely the product of *in situ* diversification) show more specialization compared to more broadly distributed species (those taxa more likely to be initial colonists; Fig. 3). At intermediate modularity and nestedness, the distributions of the number of links assigned to each hemipteran species showed the least deviation from the METE prediction (Fig. 3), suggesting that at the transition from primary succession to evolutionary assembly, these plant-herbivore communities reach statistical steady state.

## Development of genetic discontinuity at different trophic levels

The analysis of available genetic data presented here indicates 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 the sap-feeding herbivores in this study showing structure at smaller scales compared to detritivorous crickets and predatory spiders (Table 1, Fig. 2). Population structure within species allows for populations to take independent evolutionary trajectories, especially when aided by other evolutionary processes that may be acting differentially across each species’ range. 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), and sexual signaling (Mendelson & Shaw, 2005; Percy & Kennedy, 2006; Magnacca *et al.*, 2008; Goodman *et al.*, in review).

The sap-feeding Hemiptera group *Nesosydne* (Goodman *et al.*, 2012) provide evidence that some period of geographic isolation preceded divergence of sexual signals (Goodman *et al.*, in review). Shifts in plant host use are also involved in the process of diversification in this group (Roderick & Percy, 2008). In a similar radiation of leafhoppers, *Nesophrosyne* (Bennett & O’Grady, 2013), host plant specialization was implicated in driving species radiations up until approximately 1 million years ago, when plant niches were mostly exhausted on Maui; following this period, speciation, largely on the Hawaii Island, shifted to geographic mechanisms of diversification. Our network analysis indicates that specialization and modularity begin to show pronounced signals in network data on Maui (Figs. 4, 3), in agreement with the *Nesophrosyne* results and indicating 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. Other taxa at lower 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 & Kennedy, 2006; Mendelson & Shaw, 2005; Magnacca *et al.*, 2008; O’Grady *et al.*, 2011). As a contrast, spiders, which are predatory, only develop genetic discontinuities at larger spatial and temporal scales. Most important in the context of community assembly is that endemic sap-feeding herbivores developed structure quickly - on the order of less than 0.1 million years (assuming timing of speciation is reflected by divergence in mtDNA; Table 1), with predatory spiders showing local endemicity more slowly (Table 1).

## Macroecological metrics: Network structure and steady state

Across the Hawaiian archipelago, nestedness appears to decrease non-monotonically with site age, and is highest on the geologically youngest volcano, Kilauea. Modularity, on the other hand, appears to be highest on the oldest sites, but relatively low on the two youngest sites. The high nestedness on Kilauea, is likely to result if new species arriving by immigration 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 maximum entropy. The observed deviations from maximum entropy at Kilauea are largely driven by a surplus of singleton links (Fig. 3), which may reflect a state of “incomplete” assembly, and thus lower species richness, of the plant and herbivore biotas.

Conversely, Kohala shows a statistically significant agreement with maximum entropy perhaps because the Kohala site, at intermediate age (150 ky), has had sufficient time to undergo ecological succession but is still too young to be driven away from statistical steady state by specialization and rapid in situ diversification driven by host plant preference

Interestingly, the older Maui and Kauai sites show strong deviations from expectations of maximum entropy theory (Fig. 4), which is consistent with our hypothesis that the influence of evolutionary assembly on these biotas drives them away from statistical steady state. The application of maximum entropy to ecology does not currently take into account evolution (Harte, 2011). Indeed, the use of maximum entropy in ecology is inspired by its application to physical systems whose change through time is simple and does the account for the evolutionary history of biological systems, potentially a far cry from the complex change through time produced by speciation, extinction and adaption to novel ecosystems. One possible explanation for the strong deviations from expectations of maximum entropy, compared with our middle-aged site, is that while the ages of Maui and Kauai are sufficient for evolutionary assembly driven by specialization and diversification on host plants, the advanced island ontogeny of these islands may have led to range contractions and possibly also extinction of plant species on the oldest island of Kauai (Whittaker et al 2008).

In addition, we find decreased nestedness and increased modularity on Maui and Kauai. Modularity is known to result from coevolution selectively driving the traits of interacting species towards convergence (Donatti *et al.*, 2011; Nuismer *et al.*, 2013). However, it is worth noting that the influence of certain network properties such as nestedness on stability is still equivocal, and theoretical predictions of how network properties should change over evolutionary time is lacking. Theoretical and empirical studies have either suggested that nestedness promote stability or does not (Allesina & Tang 2011; but see Suweis et al 2014). Furthermore, almost all studies of food webs have been focused primarily on single snapshots in time or snapshots over short ecological time spans of network development that do not span sufficient amounts of evolutionary time (e.g., Albrecht et al 2010). Food webs are dynamic emergent entities and its broad topological characteristics 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 Hawaii may provide ideal systems for resolving longstanding debates on the causes and consequences of network properties such as nestedness.

Furthermore, we find that single-island endemics are always more specialized than multiple-island endemics. Dietary breadth has been positively associated with geographic range size (see Slatyer et al 2013 and reference therein), but the direction of casuality is unclear (Slatyer et al 2013). Dietary breadth may be permitting the colonization of 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; Fig 3), while at the youngest site Kilauea, specialist single-island endemics are limited by low plant diversity and thus show more apparent specialization (Fig 3). Conversely at the oldest site on Kauai, where plant diversity is not limiting (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 apparent generalization but this pattern disappears when analyzing the data at the resolution of plant genera, thus suggesting that Hemiptera species endemic to Maui are no more generalized on plant genera but instead may benefit from the diversification of plant species within genera on Maui.

## Future Research

The analyses presented here indicate strong patterns of a dynamic assembly process leading to contrasting hypotheses concerning the relative importance of ecological and evolutionary depending on the evolutionary age of the community under observation. In future work we will tackle these hypotheses using detailed quantitative ecological and genomic data collected from across the Hawaiian archipelago.

1. In younger communities we hypothesize that
   1. during periods of ecological assembly, communities strongly influenced by immigration will resemble random samples from regional source pools and thus metrics describing these communities will largely match expectations of statistical steady state after primary succession has completed;
   2. the exception will be communities still undergoing primary succession, which will change rapidly through time and represent non–random samples of source pools;
   3. 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).
2. Following the same logic, in older communities we hypothesize that
   1. during periods of evolutionary assembly, if the rates of processes such as niche exploration, adaptation and speciation can be modeled and measured and are found to keep pace with immigration, the resultant communities could be driven into alternate evolutionary states that fail to meet the predictions of purely statistical theories that do not account for evolutionary dynamics (Harte, 2011);
   2. networks in such communities should exhibit higher levels of specialization and modularity (Bascompte & Jordano, 2007; Donatti *et al.*, 2011; Nuismer *et al.*, 2013).
3. Systems undergoing rapid ecological and evolutionary change are generally expected to deviate most from statistical steady state; thus we expected populations in such communities to show genetic signatures of rapid change, from bottlenecks or population expansion to selection.

### Evolutionary data: Diversification within species

The current study demonstrates that taxa of 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 is aimed at gathering genomic SNP data for focal taxa within this system that represent different trophic levels. We will use it to understand taxonomic differences in the rate of differentiation, to assess the roles of genetic fusion and fission (Gillespie and Roderick 2014), and to detail the relative rates of speciation and extinction across the island chronosequence. When coupled with ecological data on abundances and interactions, the population genetic data will allow us to explore explicit aspects of the Geographic Mosaic Theory of evolution, such as trait remixing and selection mosaics. (Thompson et al. 2013).

### Ecological data: Assembly of species into communities

Our results show that the island chronosequence can reveal that ecological processes dominate in younger environments, with evolutionary processes becoming more important later. However, in order to build a more rigorous understand of the assembly process future work is focusing on conducting broad sampling of all macroscopic arthropod taxa at a large number of sites across the age gradient, thus allowing 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).

The current study provides a framework for the use of chronologically arranged oceanic island systems to examine the interplay between evolutionary and ecological processes in shaping biodiversity and demonstrates how this approach can provide insights into the development of communities over ecological–evolutionary time, and the dynamic feedbacks involved in assembly.

# Acknowledgement

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) for arthropod lineages found within the study sites. Where molecular dating analyses 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 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. 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 (distribution of number of links assigned to each herbivore) across sites and different biogeographic classifications of taxa. Top panels show that networks deviate most from MaxEnt 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 conform to maximum entropy; thus the observed minor deviation on Kohala is not different than expected by chance. Kohala shows minimal modularity, and maximal connectance. The bottom panel shows the number of links for island endemics versus multi-island endemics. Endemics show lower linkage overall, but significantly increase on the middle aged site Maui (highlighted with dotted box). Kohala shows increased linkage overall (highlighted with solid box). When looking at links to plant genera this pattern holds except that endemics on Maui no long show a difference in generality, indicating that the pattern is driven in part by plant diversity.

**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. While the sign of the z-score depends on null model and method of calculating modules (see supplemental figure) the overall trend is robust. Some level of nestedness is likely a statistical property of these networks, however it could also be driven by stabilizing ecological mechanisms. Modularity is thought to be a sign of coevolution driving convergence in traits of plants and herbivores. Note the very interesting peaks on Maui where adaptive diversification may be at its maximum.

# References

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. & others (2014) Herbivores and nutrients control grassland plant diversity via light limitation. *Nature*.

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.

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. (2010) Island biogeography of remote archipelagoes. *The theory of island biogeography revisited*, 358–387.

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.

Harte, J. (2011) *The maximum entropy theory of ecology*, Oxford University Press.

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

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

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.

Mendelson, T.C., Siegel, A.M. & Shaw, K.L. (2004) Testing geographical pathways of speciation in a recent island radiation. *Molecular ecology*, **13**, 3787–3796.

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

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

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.

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

Ricklefs, R.E. (1987) Community diversity: relative roles of local and regional processes. *Science*, **235**, 167–171.

Rius, M. & Darling, J.A. (2014) How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology & Evolution*, **29**, 233–242.

Roderick, G., Croucher, P., Vandergast, A. & Gillespie, R. (2012) Species differentiation on a dynamic landscape: shifts in metapopulation and genetic structure using the chronology of the Hawaiian archipelago. *Evolutionary Biology*, **32**, 192–206.

Roderick, G.K. & Percy, D. (2008) *Host plant use, diversification, and coevolution: Insights from remote oceanic islands*. *Specialization, speciation, and radiation. evolutionary biology of herbivorous insects* (ed. by K. Tilmon), 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.

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.

Slatkin, M. (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**, 787–792.

Thompson, J. N., C. Schwind, P. R. Guimarães, Jr., and 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.

Wagner, W. & Funk, V. (1995) *Hawaiian biogeography evolution on a hot spot archipelago*, Smithsonian Institution Press, Washington, DC.

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.

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