Response to reviewer comments

We thank Prof. Ricklefs and the two anonymous reviewers for their detailed comments that have enormously improved the clarity and quality of our manuscript. Below we detail our response to their input.

Comments from handling editor, Prof. Robert Ricklefs

- Part of the problem is that many of the statistical techniques, particularly with respect to network theory, are insufficiently explained for most readers. The structure of the sampling regime is not adequately described, and it would appear to be not strictly comparable between the islands.

  We have substantially re-written the methods and results sections to better detail statistical techniques and make more clear how focal sites were selected and plant-herbivore interaction data were compiled.

- A major problem from my perspective is that the "hypotheses" refer to qualitative expectations of patterns and are not closely linked to underlying mechanisms. Thus, the "tests" of these hypotheses do not discriminate between well–stated alternative mechanisms, but relate more to generalized expectations based loosely on ecological and evolutionary theory.

  We have re-framed the paper to focus first on discovering new patterns in the context of the Hawaiian chronosequence using ecological theory as a guide to identify potentially interesting biological outliers. We then use these patterns to generate new hypotheses about the feedback between ecological and evolutionary mechanisms in community assembly

Line 34 it is not clear that HI is beyond equilibrium dynamics; perhaps the steady state is very low; if this has a different meaning, then it should be made more explicit.

  This has been changed to make no claim about equilibrium

Line 36 by ecological dynamics, do you mean long term environmental change caused by post-eruption succession? This should be explicit in the abstract.

  We have made more clear that by "ecological dynamics" we are referring to those mechanisms such as biotic filtering and demography more often associated with ecological time scales.

Line 52 presumably you are referring to species within clades of organisms at lower trophic levels; the trophic level itself doesn't develop divergence. The same comment applies to the next sentence.

  Changed accordingly

Lines 142-3 predictive power doesn't come from patterns

  This phrasing has been removed

Lines 169 ff. these predictions appear to be ad hoc, and not based on an underlying set of mechanisms. Why, for example, would primary succession necessarily override the consequences of non-random colonization from the mainland? What is it about the process of succession that might produce this effect, and how would one test for it in comparison with the outcome of alternatives? The basic problem is to use patterns to distinguish between underlying processes, which make unique predictions. Otherwise, this becomes an exercise in describing pattern rather than using pattern to distinguish between hypotheses (useful in its own right, of course).

Further analysis and more detailed data are certainly needed to use pattern to distinguish process. In recognition of that fact we have refocused our paper to present patterns and use preliminary analyses to generate more meaningful hypotheses that can be tested with data that are fast accumulating about the biota of volcanic archipelagos.

Lines 213 ff. it might be good to state explicitly that you are going to limit your comparisons to ohia forests at mid elevation, if that is in fact how you are going to approach the problem.

We have made this more explicit in the methods section

Lines 245 ff. at what level is genetic variation relevant to the potential for divergence? Variation might be categorized within a locality, within a volcano, and between volcanoes on the same island; Fst values are mostly based on neutral variation, and perhaps are not so relevant to differentiation by selection. How will you use Fst values to estimate the potential for evolutionary differentiation and perhaps species diversification within islands?

We have made more explicit that for evolution of novel variants to arise gene flow must be reduced; F_{ST} captures the spatial structure that comes from reduced gene flow.

Lines 251 ff. it is not explicitly stated whether data exist on the distribution of herbivores across plant species within each site. How exactly were the bipartite networks constructed?

We have further clarified this in the methods section, making clear that site-specific data do not exist for plant host use so instead we make the reasonable assumption that if a known host plant is present at a site it will eventually be used and thus included in the network.

Lines 269 ff. this description will be understood only by individuals who work on network analysis.

We have further explained nestedness and modularity in the methods section for non-specialist readers

Line 290 ff. this hypothesis is more an assumption, in the sense that it is not based on an underlying process regarding how evolution should lead to specialization over time. The specialist-generalist trade-off does not appear to evolve in a single direction, and

different levels of specialization might be optimal depending on all sorts of considerations. Thus, "testing the hypothesis" is really more like 'describing the pattern', since theory concerning the specialization-generalization spectrum is not taken up here. The same comment applies to the next paragraph, and really concerns how we understand the word "hypothesis."

This hypothesis has been removed

Line 308 it strikes me as odd that the variation among sites within volcanoes is greater than the variation between volcanoes-or am I interpreting these values incorrectly? It also seems from Figure 1 that the distribution of sampling sites on the three islands is quite different, which would seem to confound space and age.

We agree that this result is unexpected. Rather than the more typical isolation by distance scenario such a partitioning of molecular variance could instead be consistent with, for example, a process of repeated colonization events between volcanoes leading to most molecular variance being present on many volcanoes while little variance is explained by unique variation between volcanoes. The primary difference in sampling between Maui and Big Island is that genetic data are only available from one volcano on Maui thus between volcano comparisons are not possible.

Line 324 some of the Hawaii species apparently have divergence times older than the island, i.e., 1.15 Ma > 0.5 Ma. How are we to interpret genetic variation within the species that are older than the island, as much of the genomic variation in the ancestral population is carried by a small number of colonists; or are the older age estimates the consequence of extinction or limiting sampling?

Bias of some sort is aways possible, or an incorrect calibration. It could also be that alleles that get to the big island (even in one individual) are older than the big island (hopping down island chain), and if those alleles get separated in to different populations, and then diverge to species, the genetic variation observed between groups could be older than the habitat.

Line 339 what exactly does the idea of maximum entropy represent, and what is its status as a null model. Since this is an important point, a little more explanation would be helpful to the uninitiated reader.

We have greatly expanded the conceptual motivation for maximum entropy, highlighting its interpretation in the context of ecology and clarifying that it serves to provide a yard stick against which to compare real assemblages. If those assemblages fail to meet the predictions of maximum entropy then there is evidence for the system to be out of statistical equilibrium and thus warrant further exploration of the mechanisms that could lead to that deviation.

Lines 345 ff. this is beginning to sound a little ad hoc, in terms of explanations for patterns.

We have provided more motivation for analyses in the methods section

Line 368 can you make a concluding sentence that describes these dynamics, or at least the change in patterns from youngest to oldest?

> *This phrasing no longer exists in the new submission*

Lines 391 ff. this is an extremely interesting statement that would benefit from additional explanation.

> *The statement of interest comes from the work of Bennett & O'Grady (2013). We have made this source more clear and contextualized it more within our own results.*

- We have also attended to the following more minor corrections:

  Lines 20-22 these superscript numbers do not appear in the author list.

  Line 59 how about "Our analyses suggest"

  Line 94 italicize Anolis

  Line 415 "suggest"

  Line 488 "assembly"

  Line 542 "REFERENCES"

Comments from Reviewer 1

- The genetic data and network analyses don't seem to hang together to make a clear coherent story. Instead they seem merely suggestive that older communities are structured by evolutionary processes, and it is not always clear that other explanations for the patterns observed can be ruled out. Overall, the paper has a "preliminary" feel to it and I believe it would be much improved if the intentions specified in the "Future research" section of the paper were carried out.

> *Similarly to comments from Prof. Ricklefs, we have addressed this by re-focusing the paper to more explicitly address how the model system of Hawaiian arthropods can be used to address eco-evolutionary questions. We then present our analysis of mostly published data as a preliminary proof-of-concept.*

- There is quite a bit of literature around these days looking at ecological vs. evolutionary processes as drivers of community assembly. I suggest the authors look at the following to at least provide some deeper background to their thinking about this question.

> *We have now included a discussion of the literature and how our study fits within the broader context. We have incorporated as much of this new literature as possible, though are constrained by citation limitations.*

Lines 169 and 178 More explanation is needed for the two hypotheses. For hypothesis 1, why should early assembly appear random? Are there other circumstances that would produce random assembly? For hypothesis 2 more explanation is needed for why

evolutionary processes should result in higher levels of specialization. I would think that this might depend upon the type of selection regime. Can species sorting through time also result in specialization?

These hypotheses no longer exist in the new submission.

- I am not sufficiently knowledgeable to judge the molecular techniques or divergence dating analysis. However I note that mtDNA have been used in at least one case to look at divergence times. Coyne has found that if closely related species hybridize, then mtDNA cannot be used for this purpose because the mitochondria of one species can be completely captured by the other, thus obscuring divergence. The authors need to show that this is not a problem for their analysis.

  It is possible that these species are older than we inferred from differences in their mtDNA (if similar DNA gets into 2 species it takes longer for the species to show DNA difference), but certainly the data show that some divergence between species is of the time scale noted, that is, quite recent. We now mention this issue (end of third paragraph of Discussion).

- Generally, more explanation of network analysis and how it is used and what can be inferred from it is needed. Ditto for maximizing information entropy. The authors cite other papers that use these methods, but these methods should be explained clearly in the supplement to this paper so they can be easily evaluated.

  We have expanded the methods and results section to better explain the origin and meaning of the results. We have also included content in the introduction to better motivate conceptually the use of maximum entropy in our analysis.

Line 213 I would use the term "replicate climatic and biotic templates" rather than "replicate communities" here since the communities you are studying have different assembly and evolutionary histories.

This wording has been removed

Line 228 and elsewhere. accession numbers not given

Accession numbers will be procured upon acceptance for publication

Line 278 What aggregate properties were maintained in the null models and how might they affect you inferences. Some info is given in the supplement but more detail is needed.

We have further clarified the null models in the methods section

Line 282 probabilistic and degree distribution models need more explanation, again in the supplement. Degree has a jargony character. It is used several places in the paper and it is not always clear what it means.

We have further clarified the meaning of the degree distribution, its derivation, use and interpretation, in the introduction and methods section

Line 322 I didn't see anywhere else in the paper a discussion of this phylogenetic data. Is it different from the genetic data?

We have clarified that these phylogenetic data are from relevant publications but not the genetic data we analyze.

Lines 340-352 Patterns here are difficult to interpret (at least for me)

We have added further discussion of the generalized linear model results in the discussion section

Comments from Reviewer 2

- I think the manuscript do a poor job in connecting the results reported with the vast literature on how ecological and evolutionary processes shape the geographical variation of species interactions. For example, it is safe to argue that the current theoretical framework to understand the evolutionary outcomes of species interactions is the Theory of the Geographic Mosaic of Coevolution introduced by John N. Thompson and developed by multiple research groups across the world. Accordingly, observational work, lab experiments and a variety of mathematical and computational models have been used to explore the role of species interactions on diversification. It is primal to a paper that aims to contribute to the integration of metacommunity ecology and evolution to do not ignore the current status of the literature and to actually show in which ways the paper provides news insights and contribute to the available theory (ies).

  The Geographic Mosaic Theory constitutes an evolutionary hypothesis of selection mosaics, coevolutionary hot and cold spots, and trait remixing by dispersal; it aims to explain how coevolution reshapes interactions across different spatial and temporal scales, and analyzes the role of species interactions on diversification. Exploring the geographic mosaic of coevolution in the context of island chronologies would indeed present extremely interesting opportunities to further understand the micro-evolution of species interactions. Geographic mosaic theory has been developed largely in isolation from the network theoretic approach, a point acknowledged by Thompson himself in collaboration with Guimaraes and others (Guimaraes et al. 2006 Proc. R. Soc. B.) and thus integrating the two is a worthy task, but beyond the scope of our paper. Our approach is more in line with the network theoretic work, focusing on how the structure of interactions of a whole community changes over time and space, rather than on diversification of traits. However, the Geographic Mosaic theory, and in particular the role of trait remixing and selection mosaics, will be one of the key areas we will explore as we accumulate population genetic data across the chronosequence. We have included this point in the Future Research.

- Some of the key predictions tested in this manuscript are related on how evolutionary processes should shape network structure. Unfortunately, these predictions are based

on assumptions that do not hold. For example, we have now multiple theoretical evidence suggesting nestedness would be favored by natural selection (the work by Suweiss), differential extinction of species (the work by Fontaine and Thebault) and processes minimizing competition (the work by Bastolla). Therefore, to assume non-evolutionary processes would generate nestedness is not supported for our current theory. Moreover, to assume that evolution would necessarily favor specialization, and consequently, modularity is not supported by the evidence provided by both empirical and theoretical work on the evolution of species interactions in which a consumer can attack multiple prey species. Thus, the authors need to revisit carefully their assumptions and hypotheses to see if their results can still be presented as evidence of increased role of evolutionary processes in older islands or if alternative explanations are more likely to explain the observed pattern. In my opinion, and ignoring my concerns related to the data used (see below), I think that the results may provide evidence for particular outcomes of the evolutionary process in interactions among plants and herbivores.

In line with further comments from Reviewer 2 and from Prof. Ricklefs, we do not yet have sufficient data to use pattern to distinguish between multiple competing hypotheses about process. Thus instead of setting out to test hypotheses we use existing data and a combination of analytical techniques to generate a set of more meaningful hypotheses that will soon be testable with mounting ecological and evolutionary data from island systems. However, we have to disagree that theory strictly supports a deep time evolutionary cause of nestedness. Indeed the work of Suweis, claiming an abundance optimization criterion, is in conflict with the stability optimization arguments of Bastolla, Fontaine and Thebault. The work of Suweis is also consistent with previous work, which we cite, that nestedness is a likely outcome of random assembly as the crux of Suweiss argument is that increased abundance, regardless of origin, increases nestedness. A primary goal of our paper is to provide a framework for explicit tests of these such ideas.

- The use of maximum entropy theory in the manuscript is hermetic. The authors mention in multiple parts of the text the notion of the statistical steady state of ecological systems and how this steady state provides a theoretical benchmark. Nevertheless, there is no mention on how these expected distributions look like. For example, what is the degree distribution predicted by the maximum entropy theory? Even if Harte (2011) provides the expected distributions this missing information does not allow a reader not deeply familiarized maximum entropy theory to understand the results reports and the usefulness of maximum entropy theory for this work.

  This is a very valid point, and one we hope has now been addressed: in the introduction and methods sections we have greatly expanded our description and conceptual motivation for using maximum entropy.

- Because most of the available evidence for macroecological studies is correlative, it is usual to control for multiple confounding variables that may affect the reported pat-

terns, leading to spurious correlations. Indeed, this careful exploration sets a control for inference on spatial and temporal trends in large-scale studies in ecology. I missed this kind controlled analysis in the current manuscript and without any kind of sensitivity analysis there is no enough evidence to support - in my opinion - the view that age is the best explanation for the observed patterns. I would like to know how the observed patterns are affect by island size or climate and habitat heterogeneity. To assume the islands are climatic similar may not be enough and there are ways of performing more rigorous analysis.

We agree that such statistical control is the core analytical strength of macroecology. Data do not currently exist to make such detailed analysis possible. Here again we have re-cast our paper, acknowledging this data limitation, to focus on using existing data to generate a richer set of hypotheses to be tested on similarly richer data.

- I have two major concerns with the network analysis performed by the authors. First, I am afraid that the trends described are much less evident than stated by the authors. There is no linear or monotonic change in nestedness and modularity with island age. Indeed, there are three islands showing higher nestedness and the old one showing lower nestedness. Accordingly, there are two islands showing more modular networks and two areas showing less modular networks. Therefore, there is no strong evidence supporting the predictions that island age would favor modular and less nested networks. The authors need to review their interpretation of the results to avoid overemphasize the mentioned trends and, actually, to double-check if there is any trend. Second, it is impossible to verify in the current draft the quality of the data used to build up the ecological networks. The only information available is: "We compiled plant-herbivore networks from published sources as described in the main text. Table 1 lists publications used in compiling these networks" (Supplement). Nevertheless the Table 1 of the Supplement is not available for reviewing ("As part of the final submission we will make available our compiled list of Hemiptera (a typo here) and their plant hosts from published sources"). As a consequence, it is impossible to me, as a reviewer, to make any evaluation on how adequate is the dataset used to build up the networks and how comparable the networks are. Thus, the central information for this manuscript is missing and I am afraid I cannot just believe that the data are adequate for the analysis performed. Along the same lines, it is not clear to me if the interactions were actually observed in each island or just assumed because there is information in the literature that a given insect feeds on a given co-occurring plant elsewhere. If the latter case is true, all the analysis is invalidated as a test of local, evolutionary assembly of ecological networks since it will be based on the assumption that there is no local specialization.

In line with similar concerns from Prof. Ricklefs we have re-cast the manuscript as presenting a framework for integrating ecological and evolutionary perspectives on biodiversity using island chronosequecnes and ecological theory to do so. We have also made network construction and analysis more transparent in the methods. From this it should be clear that

indeed interactions are derived from the spatial coincidence of herbivores and their potential host plants. This undeniably ignores the possibility of local specialization; however, local specialization is unlikely due to the already restricted (i.e. localized) ranges from which species descriptions are based and because the taxonomic level of specialization is typically that of genera.

Abstract the way network analysis is mentioned in the methods do not make clear why the authors performed this set of analysis.

We now make clear that these metrics summarize network structure.

Line 70 To state that "genomic and computational techniques have re-opened scientists' eyes to the rich and dynamic interplay between these processes" (ecology and evolution) is unfair with the work of multiple research groups that have been using multiple approaches, from naturalistic observations to detailed lab experiments, to explore the interplay of ecology and evolution.

We have removed this language

Line 84 I missed some discussion on the limitations of using islands as study case.

We have generally clarified our discussion of oceanic archipelagos as model systems. While the advantages of island systems are largely universal (e.g. their biotic simplification, discretization in space and time) their limitations are system specific and not unlike limitations possibly encountered in any mainland systems (habitat degradation, invasive species, etc.). Thus we do not address these simultaneously case-specific and yet non-unique limitations.

Line 106 This is a quite limited definition of ecology and I would doubt that ecologists that are not community ecologists would agree.

We have removed reference to "ecology" and instead refer to "ecological mechanisms" such competition and neutral drift.

Line 116 The use of network theory is much older than suggested in this sentence. See the work by Cohen, Margalef, Odum, Paine, May, and Pimm and the vast literature on food webs on the 70s and 80s.

While we acknowledge this earlier work, the more relevant application of network theory in ecology to our work comes more recently as indicated by our citations.

Line 133 This paragraph would be benefited for a synthesis of what we know about the integration of the mentioned processes.

We have substantially re-worked the introduction to now include more context.

Line 282 Please provide the actual equations describing the probability of two species inter-act in each null model. This information should be available at least in the supplement.

We have provided a more detailed verbal description of these null models and full detailes are available in the works cited.

Line 283 It is odd to call a species an "island cosmopolitan."

We retain this term as "cosmopolitan" refers to widely distributed species and here we wish to specifically refer to those species widely distributed across the islands, but not globally.

Line 458 Specialization is not evidence for coevolution. Coevolution can favor generalization and specialization can emergence in the absence of coevolution.

Language conflating specialization and coevolution has been removed

Lines 487 to 491 This part is speculative and based in a non-published paper that is not available to the audience yet. I suggest removing it. *The section has been removed.*

Line 488 A typo: assembly.

Corrected