Supplement to "Community assembly on isolated islands: Macroecology meets evolution"

A. J. Rominger[1*], K. R. Goodman[1*], J. Y. Lim[2*], F. S. Valdovinos[3*], E. Armstrong[1], G. M Bennett[4], M. S. Brewer[1], D. D. Cotoras[2], C. P. Ewing[5], J. Harte[1], N. Martinez[3], P. OGrady[1], D. Percy[6], D. Price[5], G. K. Roderick[1], K. Shaw[7], D. S. Gruner[8#], R. G. Gillespie[1#]

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. Integrative Biology, University of Texas, Austin, Texas 78712

5. Biology, University of Hawaii, Hilo, Hawaii, 96720-4091

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

* Contributed equally;

\# co-senior authors;

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

1 Time Estimation Methods

Sequence data for two mitochondrial gene regions, Cytochrome Oxidase 1 (COI) and the large ribosomal subunit (16S), from Hawaiian *Tetragnatha* spiders were obtained from existing databases and new sequences (GenBank numbers XXX-XXX). Sequences were generated following protocols described in Gillespie (2004). Divergence time estimations and topology inferences were performed using a Bayesian relaxed-clock method implemented in the BEAST v1.7.5 (Drummond & Rambaut, 2007). The times to most recent common ancestors (TMCA) were estimated for all terminals using a fixed Hawaiian *Tetragnatha* COI specific rate of molecular evolution: 3.55% per million years. This was calculated by averaging COI rates from *Tetragnatha quasimodo* and *Tetragnatha restricta*, each of which are situated with reciprocally monophyletic populations on Maui and Hawaii. To generate the most conservative rate possible, the date of the earliest emergence of the island of Hawaii was used, 0.9 million years (see discussion in Supplementary Information from (Goodman *et al.*, 2012). The rate obtained is very similar to the revised general insect COI rate of 3.54% per million years (Papadopoulou *et al.*, 2010). The separate codon positions of COI were combined under a single clock rate but were given separate models of molecular evolution. The 16S rate was estimated based on the COI rate. Appropriate models of molecular evolution were chosen using PARTITIONFINDER (Lanfear *et al.*, 2012) and are as follows: 16S (254 bp) TrN+I+G; COI (484 bp) pos1 TrN+I+G, pos2 TrN, pos3 HKY+G. The ingroup was constrained as monophyletic, and analyses were run for 50 x 106 generations with a random starting tree under a "Speciation: Birth-Death" tree prior. Trees and parameters were sampled every 1,000 generations. TRACER version 1.7.5 was used to assure the estimated sample sizes were large enough, and the resulting posterior distribution of trees were combined using TREEANNOTATOR version 1.7.5 after discarding the first ten percent of trees as burn-in. The final consensus tree is shown in Figure 1.

2   Compilation of networks and metric validation

We compiled plant-herbivore networks from published sources as described in the main text. Table 1 lists publications used in compiling these networks.

Researchers have put forward to set of "network metrics," including nestedness (Bascompte *et al.*, 2003; Ulrich *et al.*, 2009) and modularity (Newman & Girvan, 2004; Olesen *et al.*, 2007), to understand the complex structure of ecological networks. Null models are used to evaluate the statistical significance of these metrics and to compare between networks of different size Ulrich *et al.* (2009). We compare the results derived from two common null models: the "probabilistic null" of Bascompte *et al.* (2003) using the relative degree distributions of plants and herbivores as weights while randomizing links and suffers from high Type II error (Ulrich *et al.*, 2009); the "fixed-fixed null" (Ulrich *et al.*, 2009) maintains the exact number of links assigned to each species while randomizing which interactors fill the requisite set of links and suffers from high Type I error (Ulrich *et al.*, 2009). We find that using these different null models does not change any trends in our network statistics across the Hawaiian chronosequence but different null models do influence the sign and significance of the network metrics (Fig. 2). We therefore do not interpret the sing or significance of the metrics but only their relative trends across site age.

Because these networks are based on opportunistic data associated with species descriptions, and not based on standardized ecological surveys, we cannot interpret patterns in network metrics without evaluating possible sampling biases (Nielsen & Bascompte, 2007; Gibson *et al.*, 2011; Rivera-Hutinel *et al.*, 2012). To do so we rarify networks by the number of Hemiptera species included and, for each subsampled network, re-calculate nestedness and modularity z-scores. This rarefaction procedure shows that nestedness is very sensitive to network size (Fig. 4), a known property of nestedness (Nielsen & Bascompte, 2007; Gibson *et al.*, 2011; Rivera-Hutinel *et al.*, 2012). However the relative nestedness z-scores across networks remain qualitatively similar to those observed for the complete networks (Fig. 4). Modularity depends on network size in a more variable way (Fig. 4). Modularity

is expected to decrease with network size (Rivera-Hutinel *et al.*, 2012) and so the marked increase in modularity with network size on Haleakala is unexpected. However in light of the large number of highly specialized taxa this pattern is more reasonable—if most species only have within module links then removing these species through subsampling will only reduce overall modularity. Thus this pattern speaks to the high level of specialization on Haleakala, and to a lesser extent at Kokee which also shows a slight increase in modularity with network size (Fig. 4).

References

Bascompte, J., Jordano, P., Melian, C.J. & Olesen, J.M. (2003) The nested assembly of plant-animal mutualistic networks. *Proc. Natl Acad. Sci. USA*, **100**, 9383–9387.

Drummond, A.J. & Rambaut, A. (2007) Beast: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology*, **7**, 214.

Gibson, R.H., Knott, B., Eberlein, T. & Memmott, J. (2011) Sampling method influences the structure of plant-pollinator networks. *Oikos*, **120**, 822–831.

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

Goodman, K., Welter, S. & Roderick, G. (2012) Genetic divergence is decoupled from ecological diversification in the hawaiian nesosydne planthoppers. *Evolution*, **66**, 27982814.

Lanfear, R., Calcott, B., Ho, S.Y. & Guindon, S. (2012) Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, **29**, 1695–1701.

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

Nielsen, A. & Bascompte, J. (2007) Ecological networks, nestedness and sampling effort. *Journal of Ecology*, **95**, 1134–1141.

Olesen, J., Bascompte, J., Dupont, Y. & Jordano, P. (2007) The modularity of pollination networks. *Proc. Natl Acad. Sci. USA*, **104**, 19891–19896.

Papadopoulou, A., Anastasiou, I. & Vogler, A.P. (2010) Revisiting the insect mitochondrial molecular clock: the mid-aegean trench calibration. *Molecular Biology and Evolution*, **27**, 1659–1672.

Rivera-Hutinel, A., Bustamante, R., Marín, V. & Medel, R. (2012) Effects of sampling completeness on the structure of plant-pollinator networks. *Ecology*, **93**, 1593–1603.

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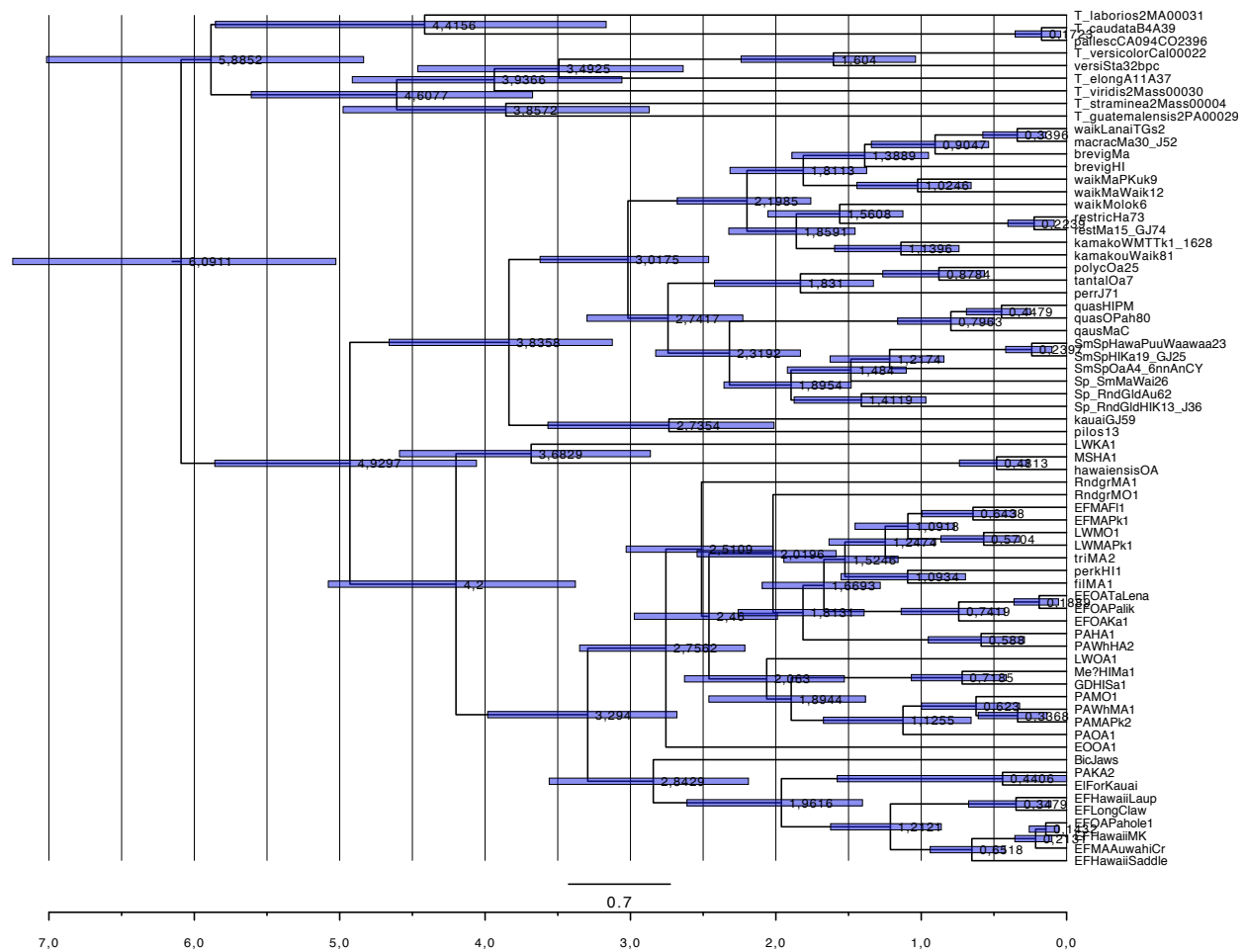
3 Supplemental Figures

Figure 1: Time callibrated phylogeny of Hawaiian *Tetragnatha* spiders estimated with BEAST (Drummond & Rambaut, 2007) used for estimating timeframe for genetic divergences.
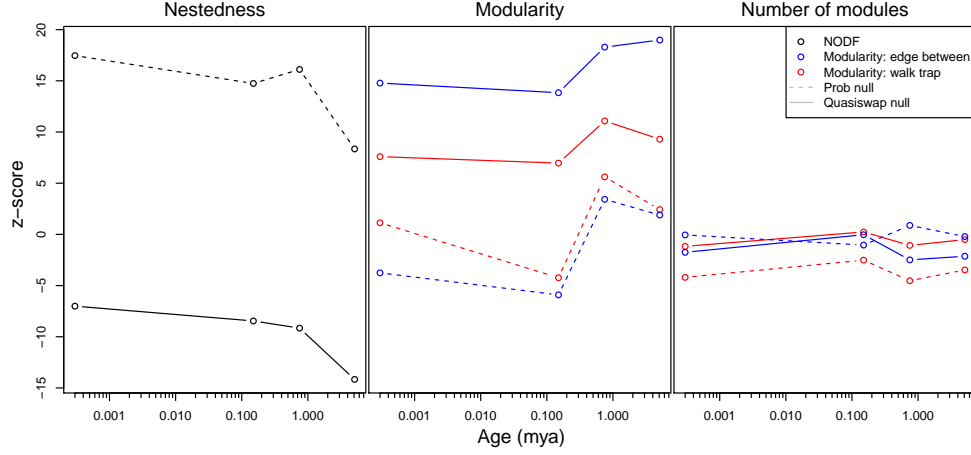
Figure 2: Comparison of different null models ("Prob" and "Quasiswap") used to standardize network metrics and comparison of different algorithms for assessing modularity ("edge between" and "walk drap"). Choice of null model has a strong influence on the sign and magnitude of metrics but not on their relative trends. The different modularity algorithms lead to largely similar qualitative patterns.
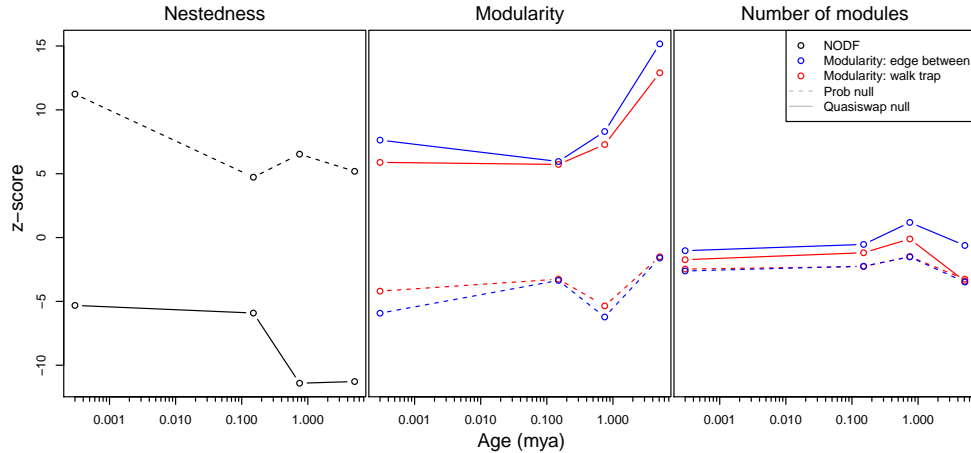

Figure 3: Metrics NODF and modularity calculated for networks based on more biogeographically conservative assignment of Hemiptera to localities. Colors and metric specifics as in Figure 2.
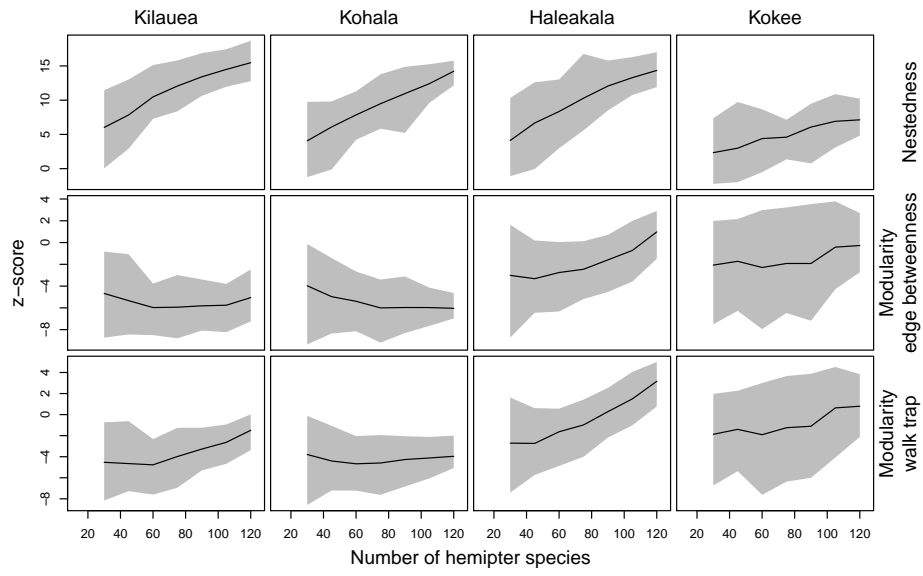
Figure 4: Result from rarification analysis showing sensitivity of network metrics to number of Hemiptera sampled.

Supplemental Tables

As part of the final submission we will make available our compiled list of Hemipitera and their plant hosts from published sources.

Table 1: Published sources of trophic information used to construct networks.