**Foundation species genotypic variation generates ecological network structure**

Matthew K. Lau1, Louis J. Lamit1, Rikke R. Naesbourg1, Stuart R. Borrett2, Stephen M. Shuster1, Catherine A. Gehring1, Miguel A. Fortuna3 and Thomas G. Whitham1

1Department of Biological Sciences and Merriam-Powell Center for Environmental Research, Northern Arizona University, Flagstaff, AZ 86011-5640, USA. 2Department of Biology and Marine Biology, University of North Carolina Wilmington, 601 South College Road, Wilmington, NC, 28403-5915, USA. 3Integrative Ecology Group, Estación Biológica de Doñana, EBD-CSIC, Sevilla, Spain.

Corresponding Author: Matthew K. Lau, Harvard Forest, 324 N Main St, Petersham, MA 01366, Email: mkl48@nau.edu, Phone: 928-523-9138

Target: Nature (Letters)

1500 words

3-4 small figures

30 references max

* Begin with fully referenced paragraph 200-300 words, aimed at readers in other disciplines = intro (2-3 sentences) + main conclusions (1 sentence, “here we show…”) + general context (2-3 sentences, how does this move the field forward)
* “References, title, author list and acknowledgements do **not** have to be included in total word counts.”
* Double space
* Contributions should be organized in the sequence: title, text, methods, references, Supplementary Information line (if any), acknowledgements, author contributions, author information (containing data deposition statement, competing interest declaration and corresponding author line), tables, figure legends.

Title

Text

Methods

References

Supplementary Information line (if any)

Acknowledgements

Author Contributions

Author Information (containing data deposition statement, competing interest declaration and corresponding author line)

Tables

Figure Legends

**Text**

Biological evolution occurs in the context of complex networks of interacting species. Ecosystems are predicted to undergo both ecological and evolutionary changes in response to environmental change, such as shifting climate. Understanding how evolutionary processes, such as natural selection, will affect the structure of ecological networks is essential for understanding ecosystem dynamics. However, no study has yet demonstrated a genetic basis to ecological network structure, which is the raw material that natural selection acts on. Here, we use both a long-term experimental common garden with genotyped individuals and a natural riparian forest of the foundation tree species, *Populus angustifolia*, to test how genetic variation contributes to the interaction network structure of a model community comprised of epiphytic lichens. We found three main results: 1) lichen communities showed significant unipartite (i.e., one mode) network structure that was similar between the common garden and natural stand, 2) genotype significantly influenced lichen species interactions within an individual tree also showing strong correlation with bark roughness, a genetically based trait known to influence bark lichen communities, and 3) bipartite (two mode) genotype-species networks, comprised of the foundation species and its associated lichen community, showed significant modular structure in both the common garden and natural stand. These results demonstrate strong support for a genetic basis to ecological network structure and the potential for selection to act in complex ecosystems. This work sets the stage for studies that address greater complexity in the evolution of biological systems and provides a framework for the discovery of evolutionarily dynamic compartments in ecosystems.

A large body of work on plant associated arthropod communities has shown that genetic variation in a foundation species can influence community composition, diversity and stability (e.g., Wimp et al. 2005, Crutsinger et al. 2006, Keith et al. 2010); and of particular relevance to plant-pollinator networks ecology, Burkle et al. (2013) found that floral visitation varied significantly among genotypes of *Solidago altissima.* In addition to arthropods, research in other systems, such as plant-plant (Lamit et al. 2011) and plant-fungal (Barbour et al. 2009, Busby et al. 2013), have also demonstrated community level impacts of genetic variation in dominant plant species.

In ecological communities, plant population dynamics depend on animal-herbivore, pollinator, and seed disperser interaction networks (Jordano et al. 2006, Schupp and Fuentes 1995, Fontaine et al. 2011). At the scale of ecosystems, trophic networks or food webs direct and control the rates of energy and nutrient flux (DeAngelis et al. 1989).

(i.e. dominant species that create locally stable conditions, *sensu* Ellison et al. 2005)

Community genetics studies have shown that genetic variation contributes to variation in interactions with dependent communities.

Although these and other studies have demonstrated the effect of genetic variation on interactions between plants and their associated species, few studies have explored more complex patterns of species interactions. One important example is Bailey et al. (2006), which showed that the interactions among species across three trophic levels depended on cottonwood (*Populus* spp.) genotype. Briefly, the study found that genotypic variation influences traits that determined resistance to a galling insect (*Pemphigus betae*). The shifts in gall abundance were then found to correlate with avian predation on galls. This study highlights the potential for indirect effects of genetic variation to propagate through networks of interacting species and trophic levels.

In order to explore how genetic variation can contribute to more complex interactions within an entire community, we used the epiphytic lichen community associated with *Populus angustifolia* (narrowleaf cottonwood), as it is comprised of species that live in close association with an individual tree and can be easily identified in the field. Previous studies in this system have shown that the composition of this community varies in response to tree genetic variation (Lamit et al. 2013). Building on these findings, we set out to determine if genetic variation in a foundation tree species can contribute to predictable variation in the interactions among associated species. To do this, we modeled species dependency networks (Araujo et al. 2011) and quantified the co-occurrence patterns on individual trees of *P. angustifolia*both in a wild stand and a common garden with trees of known genetics. We argue that this is a suitable model for detecting shifts in interactions, as interactions occur locally between individuals, patterns of co-occurrence can be used to measure shifts in the frequency with which species can interact. We hypothesized that stand level network structure can arise from three main sources of variation: 1) substrate (i.e. tree) age, 2) direct and indirect influence of variation among trees and 3) genetically based variation in tree traits. I think you need to justify a bit more why you think these 3 things matter. As heritable variation is the raw material for natural selection, a genetic basis for interaction network structure has implications for the evolution of complex interactions among species.

**Methods**

*Riparian forest and common garden stands*

Riparian forest Uintah, Utah, USA. Coordiates. Age of trees. Spacing of trees.

Common gardens. ONC. Coordinates. Planting date. Spacing of trees. Genotype sources. Pit. Coordinates. Planting date. Spacing of trees. Genotype sources.

*Lichen observations*

Quadrats. Heights and testing for height effect on composition.

*Co-occurrence Pattern and Network Analyses*

Co-occurrence analyses.

Network modeling and analyses.

*Software*

All analyses were conducted using the programming language R version 3.0.2 (R Development Core Team 2014).

*Study System and Sampling*

The study was conducted along the Weber River, UT (USA). The Weber River is cottonwood (*Populus* spp.) dominated riparian ecosystem. There are two native species *Populus angustifolia* (James) and *Populus fremontii* (S. Watson), which are both widely distributed throughout western North America and known to hybridize (Eckenwalder 1996). For the purposes of this study, only pure or highly advanced backcrosses of *P.* *angustifolia* were sampled in order to avoid the effect of the hybridization between these two species.

A stand of *P. angustifolia* near the city of Uintah, UT (GPS: N41.13903, W110.94400) was used for the wild stand survey. A total of 14 trees were chosen randomly over a 0.10 km2 area with a minimal distance of 5.56 m between trees. Only trees with *P. angustiolfia* morphology were sampled. All trees were geospatially referenced using an average of 10 points taken with a Garmin (Garmin, Ltd.).

A common garden setting was used to isolate the effect of tree genotype from the effect of the localized micro-environment associated with each individual and spatial autocorrelation. Asexually propagated clones of genotyped *P. angustifolia* individuals (Martinsen et al. 2001) were obtained from wild collections and planted randomly in a single field (0.025 km2) at the Ogden Nature Center, Ogden, UT in 1992(?). Although some replanting of replicates occurred, the trees in the garden are either the same age or at most three years younger than the original planting (T.G. Whitham, Personal Communication). A total of thirteen genotypes replicated between 3 and 8 times each, were chosen for sampling.

*Lichen Co-occurrence and Tree Trait Sampling*

To establish co-occurrence patterns among the lichen species, fine scale distribution data were collected for epiphytic lichen associated with *P. angustifolia*. Observations were made in the common garden in October 2010 and May 2011 and the wild stand in May 2012. Previous work in this system has shown that the bark lichen community is comprised of fourteen species, *Xanthomendoza galericulata, X. montana, Caloplaca holocarpa, Candelariella subdeflexa, Rinodina glauca, Lecanora hagenii, Lecanora* (unknown species)*, Phyciella melanchra, Physcia adscendens, Physcia undulata, Phaeophyscia orbicularis, Phaeophyscia ciliata, Melanelia subolivacea, Meanelia elegantula*, including both crustose and foliose lichen species that exhibit low inter-annual variation (Lamit et al. 2013). All species could be identified in the field with the exception of *Physica* and *Physciella*, due to the often small thallus size (< 1mm2) and highly similar morphology, we combined them into a generic category that we refer to here as “Phys.”

On each tree, presence or absence of each lichen species was assessed in 50 total 1 cm2 cells arrayed in a checkerboard pattern. Two adjacent 10 cm by 10 cm quadrats centered at 50 cm and 85 cm from ground level were sampled. The checkerboard sampling pattern was chosen to isolate each cell based on an average thallus size of 1 cm2. Samples were restricted to the northern aspect of the trunk to maximize the abundance of lichen and control for the effect of aspect. Thus, the thalli in each cell are expected to be spatially independent of the other cells in the quadrat, but exposed to similar micr-environmental conditions.

As previous studies by Lamit et al. (2011 & 2013) have shown that bark texture or roughness was a good predictor of lichen composition and that it has an underlying genetic basis, we chose to focus on bark roughness as a predictor of co-occurrence patterns. The bark roughness of each tree was observed within the same quadrats used to sample the co-occurrence patterns on each tree in both the wild stand and the common garden. Rough bark was defined as any disruption of the outermost layer of tissue, which in early stages of growth is typically smooth and white (Lamit et al. 2011). The presence of rough bark was visually assessed within every 1 cm2 cell (100 cm2 total). Thus, the percent of rough bark was equal to the total number of cells containing rough bark.

In the wild stand, two additional tree traits were observed: canopy cover and tree age. Percent canopy cover was quantified using a spherical densiometer (Forestry Suppliers, Inc.). Densiometer readings were taken at 1 m distance from the trunk on the northern side of each tree. Tree age data was obtained from Lamit et al. 2013. Briefly, tree cores were taken at a height of 1 m using an increment borer (Forestry Suppliers, Inc.). Annual growth rings were counted at 200X magnification using a dissection microscope (ZEISS).

*Network and Null-Model Based Co-occurrence Modeling*

Stand-level dependence network models were generated for both the wild stand and the garden using the methods of Araujo et al. (2011). This method uses conditional probability to compare the observed to the expected frequency of co-occurrences between all pairs of species. The method produces a matrix of statistically significant dependency values for each dataset. Here it is important to note that each dependency value is the probability of co-occurrence relative to the probability of occurrence across the entire sample, and therefore, it can produce a different value for each species in a pair. The result is a matrix that defines a network of significant relationships between species based on the degree to which they occur with or without another species, relative to that species occurring at all.

To explore the tree level patterns that contribute to the stand level network structure, we conducted null-model based co-occurrence analysis. We used Stone and Roberts (1991) C-score (*C-Score* = , where *r* is the total number of occurrences for a species and *S* is the number of co-occurrences of species *i* and *j* for all *n* species). A Standardized Effect Size (*SES* = ) was then calculated for each tree by standardizing the observed C-score with C-scores for co-occurrence matrices generated by permuting the matrices so that species totals were preserved (Gotelli 2002). Thus, the SES is a measure of the observed co-occurrence patterns relative to co-occurrence patterns due to chance.

*Statistical Analyses and Software*

We used a combination of univariate and multivariate statistical tests. A correlation test with Pearson’s r was used to test for the correlation between the wild and common garden networks. Linear regression and Analysis of Variance (ANOVA) were used to test for the effects of tree age, canopy cover, bark roughness and genotype on SES values. A permutation based Mantel Test was used to test for the effect of spatial distance on SES values (Mantel 1957). A vector analysis was used to assess the relationship between community composition and SES values in the common garden.

All modeling and analyses were conducted using the R statistical programming language (v3.0, R Development Core Team 2013). The methods of Araujo et al. (2011) were implemented *de novo* (see Supplementary Materials) and the *sna* package was used for network analyses and plotting (Butts et al. 2010). The *vegan* package was used for all co-occurrence modeling and multivariate analyses.

**Results**

Unipartite Network Structure

* ONC = Significant co-occurrence patterns (obs = 26.33, SES = -6.31, p = 0.0002)
* Uintah = Significant co-occurrence patterns (obs = 1.08, SES = -3.15, p = 0.002)
* ONC = Significant network patterns (Araujo et al. 2011)
* Uintah = Significant network patterns (Araujo et al. 2011)
* ONC + Uintah = significant network similarity (Mantel: r=0.51 p=0.029) \*Relativized to maxima
* Node centrality was highly correlated between networks (r = 0.7, t = 2.6135, df = 7, p = 0.035)
* Centrality was also highly correlated with total abundance for both networks (ONC: r = 0.77, t = 3.2427, df = 7, p = 0.014; Uintah: r = 0.86, t = 4.43, df = 7, p = 0.003)

Bipartite, Genotype-Species Network Structure

* ONC = Significant modularity (obs = 0.20, z = 9.64, p<0.001)
* Unitah = Significant modularity (obs=0.3243, z = 7.422347, p<0.001)

Tree Genotype Affects Lichen Network Structure Largely Driven by Roughness

* In addition to compositional affects determining co-occurrence patterns, intra-tree co-occurrence patterns are determined by genotype
* ONC = genotype affects co-occurrence patterns (F = 3.4213, num df = 12.000, denom df = 14.668, p-value = 0.01426)
* ONC = Genotype influences roughness (F = 2.953, num df = 12.000, denom df = 14.589, p-value = 0.02653)

Species accumulation curves showed that communities in the wild and the common garden were thoroughly sampled and with very similar species richness (Supplementary Materials). In the wild stand the most abundant lichen, *X. galericulata*, had an average occurrence of 57%, with the next most abundant species, *C. subdeflexa*, having an average occurrence of 10.5%. All other species in the wild stand occurred at frequencies of 2% or less.

The lichen community exhibited network structure that was similar between the wild stand and the common garden. Significant species dependency network structure was observed for both the wild stand and the common garden (Fig. 1a,b).

Dependency values that were non-zero in either network were highly correlated (r = 0.64, P << 0.0001; Fig. 1c).

The most notable difference between the two networks was the increased chance in the connections of *Lecanora* sp. (Fig. 1d).

In the wild stand, bark roughness was the primary driver of network structure. First, trees varied in significance and magnitude of their SES values, which were all negative or equal to zero (Supplementary Materials, Table 1). Second, bark roughness had a strong, significant effect on SES magnitude (R2 = 0.29. P = 0.027; Fig. 2). Third, neither tree age nor geographic distance had a significant effect on SES.

In the common garden, where the effect of environmental variation was controlled, genotype was an important factor contributing to network structure. Genotype significantly affected the SES of each tree (R2 = 0.65, P = 0.046, Fig. 3a). The SES value was also highly correlated with community composition (R2 = 0.71, P < 0.001, Fig. 3b).

Present the garden results first, since this is the bulk of the support for your main hypothesis – that genetic variation in a foundation sp can influence network structure (and you present this finding first in discussion)

**Discussion**

Genetic variation contributed to interaction networks structure with networks tending to be nested. In addition, genetic variation also contributes to the structure of interactions among associated epiphytic lichen species is directly.

*Genetic basis of species networks*

These findings support the hypothesis that genotypic variation in a foundation species can contribute to the structure of a network of interacting species. Several lines of evidence support this conclusion. First, the wild stand showed significant dependency network structure (Fig. 1a); and the genetically based tree trait, bark roughness, was a strong predictor of co-occurrence patterns (Fig. 2), while neither the effects of tree age nor geographic distance were significant. Second, the common garden network (Fig. 1b) structure showed a high degree of similarity to the wild stand network structure (Fig. 1c&d). Third, tree genotype was a significant predictor of SES values (Fig. 3a); and SES value was strongly correlated with community composition (Fig. 3b). Thus, just as numerous studies have shown that plant genotype can affect species richness, abundance, diversity, and composition, for the first time this study shown both observationally and experimentally that species interaction networks also have a genetic component.

* what happened to 1 fig per paragraph? Break these up and discuss their implications. It feels like a rehashing of the results section.
* What does it mean that genotypes have different SES values? Does a bigger negative SES mean that that genotype is better at predicting its community?
* Explain the implications for wild and garden stands having similar networks

It is important to note that although we have not directly quantified interactions between species, the dependency network structure and tree level co-occurrence patterns are good indicators of ecological interactions. Although the co-occurrence literature has a long history of dialogue about inferring interactions from spatial patterns (e.g., Diamond 1972, Connor and Simberloff 1983 and Gotelli 2002), multiple studies have been done developing and evaluating spatial pattern based network modeling methods (e.g., Kissling et al. 2012). Since organisms interact along multiple ecological dimensions; such as resource competition or mutualism, predation, or behavioral interference, and a spatial pattern based perspective departs from direct observation of a mechanism, however it contributes an integrated perspective that can compliment more mechanistic studies. We reiterate that species interactions typically occur in close spatial proximity, and therefore species co-occurrences are an indicator of the frequency of interactions.

Although this study was done with lichen, these results can be generalized to other groups of organisms, although spatial scale of interactions should be considered. First, sessile organisms associated with a host species, such as coral, barnacles, endophytic fungi, epiphytic plants and intestinal endosymbionts to name a few, are communities with high potential to exhibit similar patterns of intra- and inter-host patterns resulting from host genetics (CITE). Second, communities of species with interactions that operate at a larger spatial scale might be expected to have less response to another species’ genetic variation, as diffuse co-evolutionary theory would predict (CITE). However, taking a network perspective, previous studies have shown that weak interactions (CITE) and indirect effects (Patten and Higashi 1995) can have large, surprising effects. Give an example even.

Based on these results

* How would you predict climate change to later these interactions? What type of connections would get stronger or weaker? What about ways to assess negative impact of climate change through network structure? Some sort of network instability measurement..
  + Some sort of weighting factor, if you knew whether the interaction was positive (i.e. facilitation) or negative (i.e. competition), would be really cool.
* What about exotic species? If looking purely from a network perspective, might see more connections to be a good thing, but if it involves an exotic, not necessarily beneficial. Can you compare network structure between presence and absence of exotic species? Might expect that Tamarix, with its extensive impact on many different facets of the surrounding environment, to have a complex network structure when it is present, maybe implying that removal will have a larger effect than if there weren’t as many connections.
* What about promoting the conservation of interactions rather than just species per se?

*Conclusion*

*Evolution of interaction networks and relevance to community assembly*

In this study, we have shown how genetic variation can contribute to the structure of a network of interacting species. This has major implications for understanding how networks of species evolve, since genetic variation is the raw material for evolution. Although this has primarily been from an ecological perspective, network theory provides not only a useful tool for systems level analysis but also a common framework for bridging disciplines. These findings have implications for understanding how other networks that operate entirely or in part under principles similar to biological evolution, such as whole ecosystems.

*Observations in the wild and common garden*

Recent studies have questioned the appropriateness of common gardens to reflect community patterns in the wild (Tack refs). Our findings show that …

*Mechanisms (Bark roughness)*

- Wild = genetics, local environment and tree age

Other questions:

Do genetically similar trees support more similar networks? Genetic similarity rule

**Acknowledgments**

**References**

Robert C. Barbour, Julianne M. O'Reilly-Wapstra, David W. De Little, Gregory J. Jordan, Dorothy A. Steane, Jonathon R. Humphreys, Joseph K. Bailey, Thomas G. Whitham, and Bradley M. Potts A geographic mosaic of genetic variation within a foundation tree species and its community-level consequences Ecology 2009 90:7, 1762-1772

D. L. DeAngelis, P. J. Mulholland, A. V. Palumbo, A. D. Steinman, M. A. Huston and J. W. Elwood *Annual Review of Ecology and Systematics* Vol. 20, (1989), pp. 71-95

Eckenwalder, J.E. (1997) Chapter 1: Systematics and evolution of *Populus*. In Biology of Populus and Its Implication for Management and Conservation. Eds. R.F. Stettler, H.D. Bradshaw, Jr., P.E. Heilman, T.M. Hinckley. NRC Research Press, Ottowa, Ontario, Canada.

Fontaine, C., Guimarães, P. R., Kéfi, S., Loeuille, N., Memmott, J., van der Putten, W. H., van Veen, F. J. F. and Thébault, E. (2011), The ecological and evolutionary implications of merging different types of networks. Ecology Letters, 14: 1170–1181. doi: 10.1111/j.1461-0248.2011.01688.x

Martinsen, G. D.; T. G. Whitham; R. J. Turek and P. Keim (2001) Hybrid populations selectively filter gene introgression between species. *Evolution* 55:1325-1335.

Jordano, P., Bascompte, J. and Olesen, J.M. 2006. The ecological consequences of complex topology and nested structure in pollination webs. In: Waser, N.M. and J. Ollerton (eds.). *Specialization and generalization in plant-pollinator interactions*. University of Chicago Press, EEUU. Pages: 173-199.

Keith, A.R., Bailey, J.K. and T.G. Whitham. **2010**. A genetic basis to community repeatability and stability. *Ecology* 91:3398-3406.

Schupp, E.W., Fuentes M. 1995. Spatial patterns of seed dispersal and the unification of plant population ecology. Ecoscience 2: 267–275.

Wimp, G. M.; G. D. Martinsen; K. D. Floate; R. K. Bangert and T. G. Whitham. **2005**. Plant genetic determinants of arthropod community structure and diversity. *Evolution* 59:61-69.

**Figure Legends**

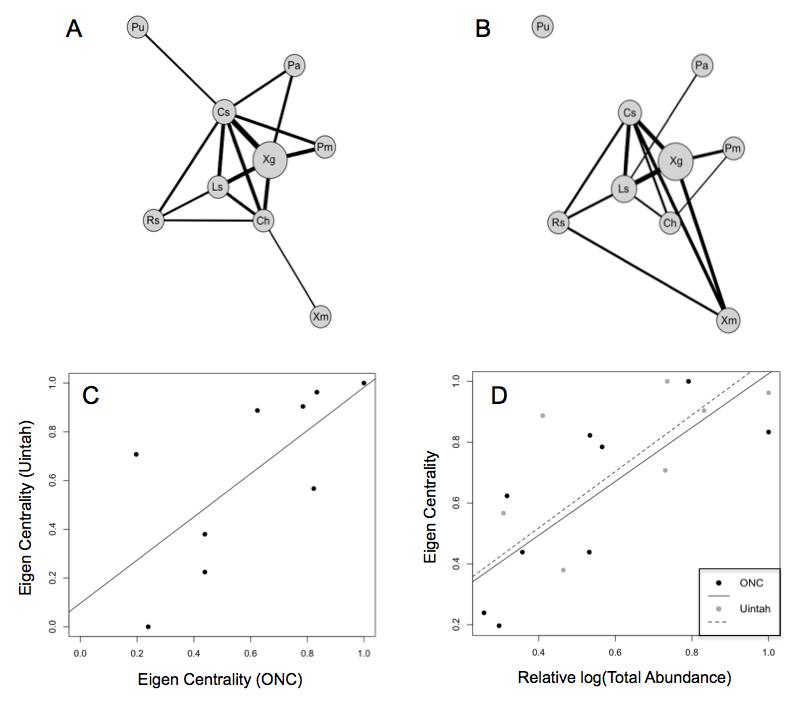
Figure 1. Significant unipartite network structure was observed for epiphytic lichens on trees of known genotype in a common garden (ONC = Ogden Nature Center, Utah, USA) (A) and individual trees in a natural stand (Uintah, Utah, USA) (B) of the foundation species, *Populus angustifolia*. Both networks are shown here with lichen species as nodes (see Methods for species names) scaled by the log of their total abundances and significant co-occurrence patterns between species shown as edges scaled by their log frequencies. The bivariate plot (C) shows the significant correlation in Eigen Centrality between the two networks. (D) The total abundance of lichen species was a significant driver of network structure for both networks.

Figure 2. Tree genotype influenced lichen co-occurrence patterns in the common garden. The barplot shows the lichen community C-Score Standardized Effect Size (SES) values for each genotype (mean ± 1SE) in order of decreasing SES. More negative SES values indicate more co-occurrences between species than expected by chance using a conservative null model that accounts for differences in both tree and lichen species totals.

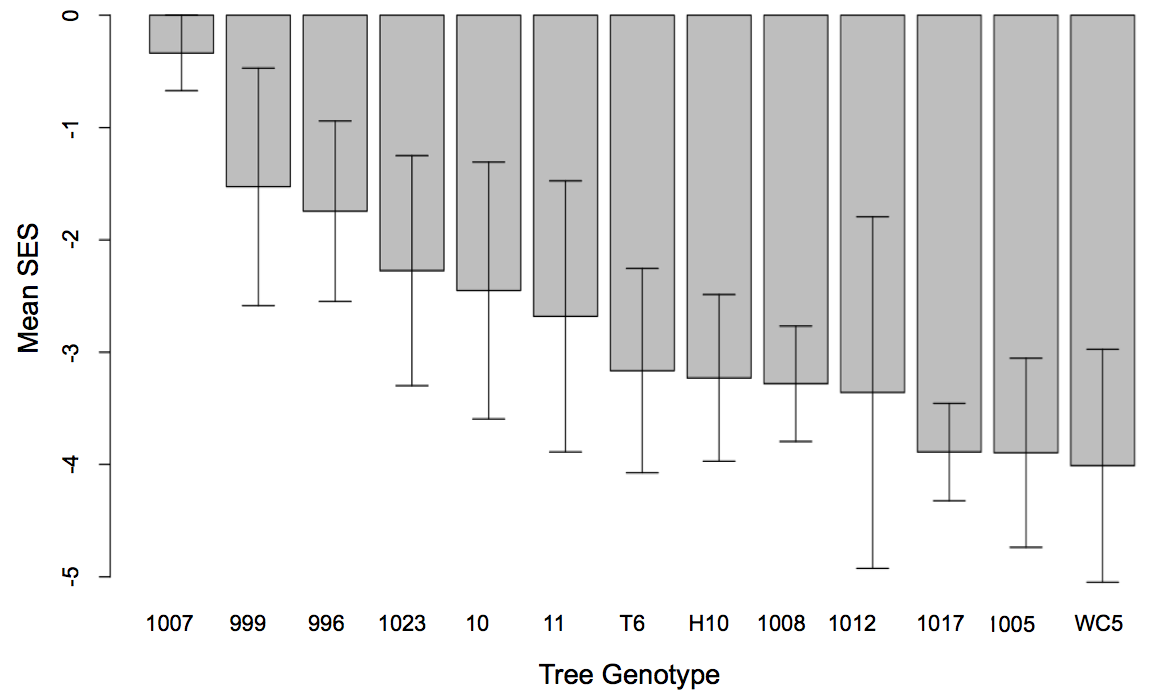
Figure 3. The lichen co-occurrence patterns were highly correlated with the genetically based phenotypic trait, bark roughness (i.e., the percentage of textured bark), in both the common garden and natural stand. The scatterplot (A) shows the mean (± 1SE) percent rough bark and SES for each genotype for trees in the common garden with SES values becoming more negative, indicating stronger co-occurrence patterns, as bark roughness increases. The lichen communities on individual trees in the Unitah natural stand (B) displayed a similar pattern with the SES values becoming increasingly more negative on trees with more rough bark.

Figure 4. Bipartite networks displayed significant modularity with modules comprised of both genotypes and species. The left most set of nodes shows tree genotypes (see Methods for genotype names) for the common garden (A) or individuals in the natural stand (B) connected to lichen species on the right. Both sets of nodes are scaled by their marginal totals (i.e., total observed individuals for tree nodes and total abundance for lichen species) and arranged by ascending totals from bottom to top. Node color shows the significant module membership for both trees and lichen species with module color having no direct relationship between the two networks, as modules were determined for each network independently.

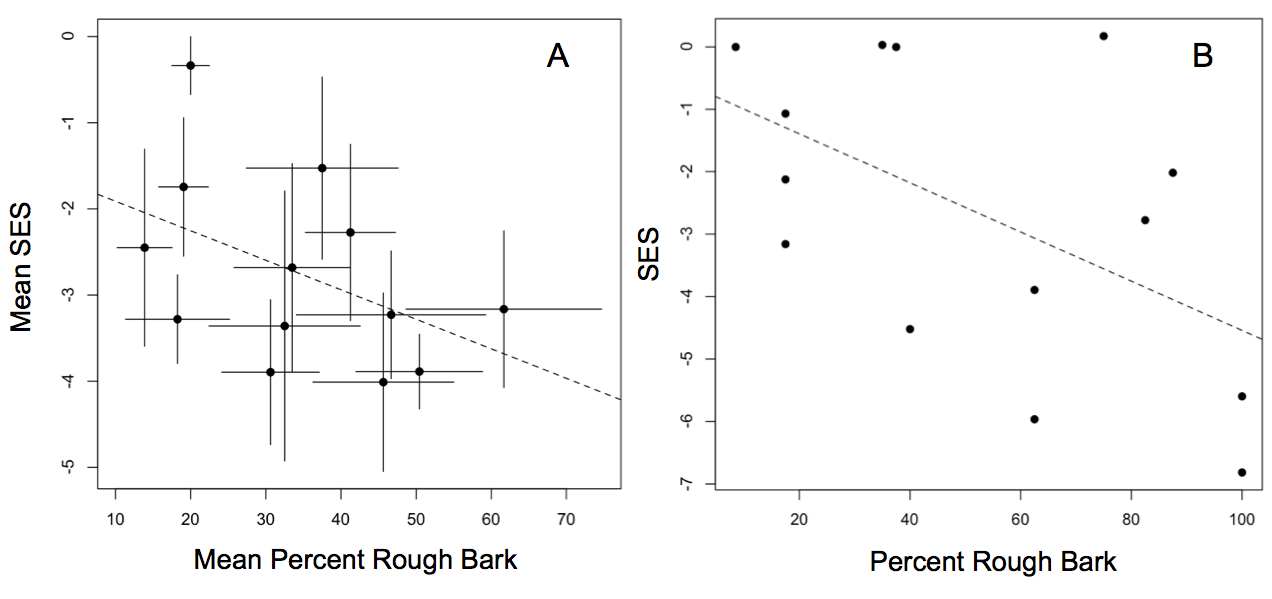
**Figure 1**.



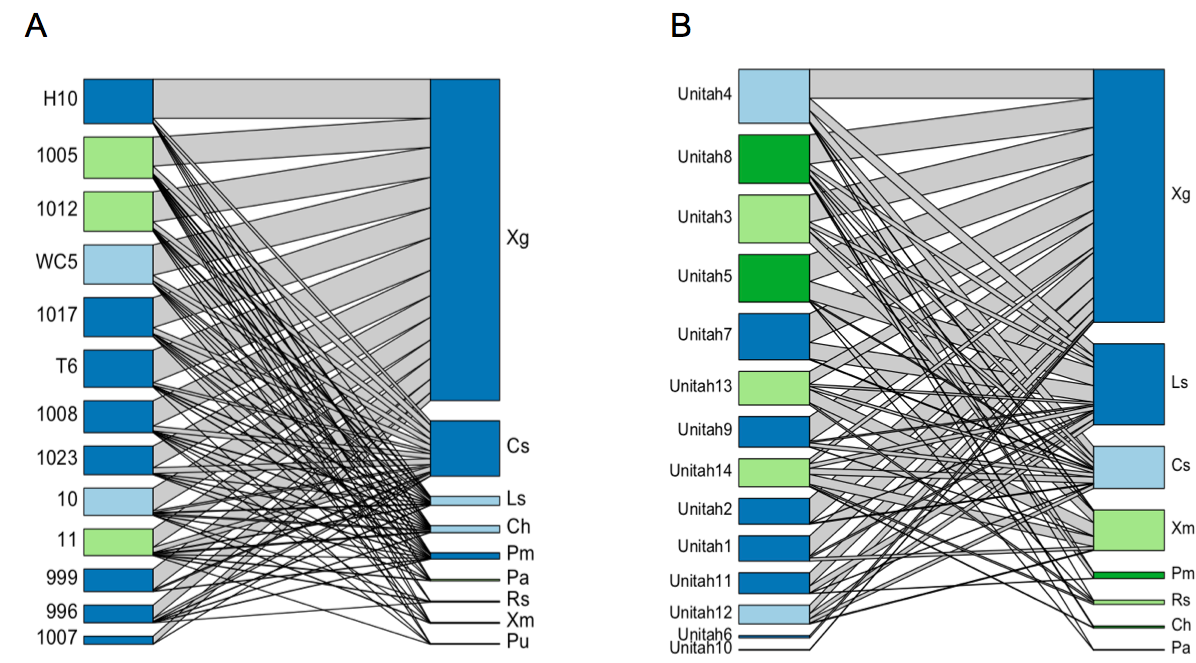
**Figure 2.**



**Figure 3.**



**Figure 4.**

****