

Species Association Networks of Tropical Trees Have a Non-neutral Structure

Michael M. Fuller
Brian J. Enquist
Andreas Wagner

SFI WORKING PAPER: 2004-06-015

SFI Working Papers contain accounts of scientific work of the author(s) and do not necessarily represent the views of the Santa Fe Institute. We accept papers intended for publication in peer-reviewed journals or proceedings volumes, but not papers that have already appeared in print. Except for papers by our external faculty, papers must be based on work done at SFI, inspired by an invited visit to or collaboration at SFI, or funded by an SFI grant.

©NOTICE: This working paper is included by permission of the contributing author(s) as a means to ensure timely distribution of the scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the author(s). It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author's copyright. These works may be reposted only with the explicit permission of the copyright holder.

www.santafe.edu



SANTA FE INSTITUTE

Species Association Networks of Tropical Trees have a Non-Neutral Structure

Michael M. Fuller^{1*}, Brian J. Enquist², Andreas Wagner¹

* Corresponding Author: (505) 277-1718
FAX (505) 277-0304
Email: mmfuller@unm.edu

1 Department of Biology
University of New Mexico
Albuquerque, NM 87131

2 Department of Ecology and Evolutionary Biology
University of Arizona
Tucson, AZ 87519

Abstract

An important question in community ecology is how species interactions influence the structure of ecological communities. Here, we ask if patterns of species associations within a diverse tropical forest in Costa Rica are consistent with neutral models of community organization or consistent with niche based models. We quantified species associations using the extent of crown overlap between trees. Crown overlap is proportional to the amount of light and soil resources shared between individuals. We studied the effects of niche structuring by representing the community as a species association network, a graph whose nodes are species. Two species are connected by an edge in this graph if the crowns of at least two individual trees in the two species overlap. In order to assess network structure predicted by neutral vs. niche structured models, we compared the structure of species association networks with those of randomized communities. Two indicators of network structure showed highly significant differences between empirically observed and randomized communities. These differences became more pronounced for large trees with a high extent of crown overlap. This indicates that factors associated with niche structuring (such as competition, facilitation, and abiotic filtering) influence the structure of the community. The effects were most striking in older, large trees. Our results also show that how we choose to represent and analyze community structure is critical for our ability to reject a neutral hypothesis.

Introduction

Community ecology aims at understanding the forces that govern the structure and dynamics of species communities. A central question in community ecology is whether niche differences among species (e.g. competition, facilitation, and abiotic preferences) influence indicators of community structure such as species abundance (Brown 1984; Rosenzweig 1995; Tilman 2000). Recently, proponents of the neutral theory of biodiversity and biogeography have taken a hitherto neglected position on this question. They argue that random demographic processes can cause observed patterns of species diversity and abundance (Bell 2001; Hubbell 2001; Hubbell and Foster 1986; Volkov et al. 2003; Volkov et al. 2004). More specifically, neutralists argue that under homogeneous environmental conditions, species have equal probabilities of establishment and survival. Random birth, death, and dispersal can thus generate communities whose structure is similar to that of natural communities. As evidence for their claims, neutralists have shown that several properties of communities, such as species relative abundance, can be reproduced by simple neutral models (Bell 2000; Hubbell 2001). Pertinent empirical evidence is equivocal. For example, Condit et al. (2002) found that the neutral theory is able to predict the composition of tropical forest plots at intermediate distances (0.2-50 km), but not at smaller and larger scales (Duivenvoorden et al. 2002).

The focus of our study is a community of tropical trees in Costa Rica (Janzen 1986) which was first analyzed by Hubbell in the 1970s (Hubbell 1979). Among other patterns characterizing this community, Hubbell found that 72% of tree species in this community had “clumped” distributions. Hubbell suggested that the non-uniform species distribution he observed could arise through the random replacement of fallen trees by new seedlings, and through a spatial autocorrelation in the random dispersal of offspring around mature trees. Hubbell’s models to explain this and similar patterns of community organization were the seeds of the neutral theory (Hubbell 2001). Here, we analyze patterns of species associations in a tropical tree community that is similar in location and composition to Hubbell’s original 1979 study plot. The dataset consists of a permanently mapped plot established by Hubbell in 1976 and surveyed by one of us (Enquist et al. 1999), using methods identical to those in Hubbell’s (1979) study.

Closely spaced trees in a forest have ample opportunity to compete. The closer two trees are to one another, the more they compete for light, space, and soil nutrients, because the extent of overlap between the branches and roots of two trees is inversely proportional to their inter-trunk distance. The magnitude of competition between neighboring trees is proportional to their similarity in size (Harper 1977). Small trees have little impact on much larger neighbors. Conversely, large trees disproportionately reduce the amount of resources available to their smaller neighbors. A neutralist would argue that such competition is uniform across the community and thus does not affect community structure. This position is encapsulated in a key assumption of neutral models for sessile organisms like trees, namely that any one individual – regardless of species – has the same probability of dying or reproducing at any given moment (Hubbell 2001).

Here, we indirectly test whether the basic tenets of the neutral theory hold for our study community. Specifically, we ask whether our study community has a non-random structure, which may indicate the presence of niche differences. To address this question, we take an approach that does not have any precedents in community ecology: we represent community structure as a graph which we call a species association network (SAN). Although graph theory

has been previously applied to the analysis of trophic relationships in food webs (e.g. Cohen and Newman 1985; Dunne et al. 2002) and studies in landscape ecology (Cantwell and Forman 1993; Urban and Keitt 2001) our approach is the first use of graph theory to quantify community structure. We use SANs to assess both empirical patterns and neutral expectations of community structure.

A graph is a mathematical object consisting of nodes and edges. The nodes in our species association network are species. An edge connects species V_1 to species V_2 if the canopy of at least one individual of species V_1 overlaps with the canopy of at least one individual of species V_2 . We characterize the structure of the species association network with various commonly used indicators of graph topology. Among them is the distribution of species with which any one species interacts; the characteristic path length of the network, that is, the average number of edges that connect any two species; and the clustering coefficient, an indicator of preferential associations among groups of species. Our motivation for using this graph-theoretic approach is twofold. First, graphs provide a simple representation for species interactions. Second, this representation is more fine-grained than other indicators of community structure, such as species abundance. It may thus have more statistical power to detect the effects of niche differences.

To assess the role of neutrality in community structure, we compare the topology of a species association network with the same network derived from a randomized community in which we reshuffled the geographical coordinates of individual trees. In doing so, we have to take into account that differences in tree spacing, body size, and canopy overlap can influence the strength of competition among trees (Harper 1977). In addition, these variables are related to one another. For example, the number of trees per unit area varies inversely with body size, which may affect the likelihood that large trees show a high extent of canopy overlap. To account for body size and canopy overlap differences, we thus studied multiple different species association networks which are distinguished by trees of different sizes and by different extents of canopy overlap between trees.

Methods

Site Location and Data Collection

We here build on earlier work by Stephen P. Hubbell (Hubbell 1979) and Brian J. Enquist (Enquist et al. 1999) who recorded the size, identity, and geographic coordinates of individual trees in an assemblage of tropical dry forest in Costa Rica. We utilize data on the spatial position of all trees found within a study plot located in the San Emilio Forest of northwest Costa Rica. The San Emilio Forest is a seasonally dry lowland tropical forest (10° 45' N, 85° 30' W) within sector Santa Rosa, Area de Conservacion, Guanacaste (ACG). Its dry season lasts five to seven months (Janzen 1986). For details on the physical and biological characteristics of the plot see Enquist et al. (1999). Janzen (1986) published a general description of the history and biotic resources of ACG.

Between 1995 and 1996, B.J. Enquist and C.A.F. Enquist mapped and identified all woody plant stems with a basal stem diameter at breast height greater than 3cm within a continuous 680m x 240m (16.32 ha) plot within the forest (Enquist et al. 1999). We will refer to this diameter simply as the *stem diameter*. The plot is composed of secondary growth forest that is surrounded by contiguous forest on three sides and bounded on one side by a narrow road. The plot is heterogeneous with respect to age, topography and degree of deciduousness. Environmental conditions such as soil moisture vary across the site. For example, the northern

part of the site contains a mesa-like escarpment that is drier than the valley located in the central part. We restricted the analysis to a 14.2 ha rectangular section of the San Emilio plot. In total, 19,873 individual plants were recorded in 1996, of which 13,639 individuals (106 species) were trees. We only included tree species in our analysis.

Interactions of Tree Pairs: Percent crown overlap

The height and diameter of a tree crown scale allometrically with stem diameter (Niklas 1994 and references therein). For trees in our study area, B.J. Enquist derived empirical allometric scaling relationships that quantify how tree height and crown diameter scale with stem diameter using tree measurements recorded in the field at San Emilio (Enquist, unpublished data). Specifically, Enquist characterized the allometric relationships between diameter at breast height, crown spread (measured as average crown radius from the trunk center), and height for most of the abundant species within the plot. Data were regressed at both the inter and intraspecific level. No significant differences in the allometric functions were observed between taxa. Here we use interspecific allometries for both canopy radius and tree height. The equations predict empirically measured values of crown height and diameter well ($R^2 = 0.918$ and 0.922 respectively). Trees within the study plot showed an average stem diameter of 12.34 cm (approximate tree height 6.4 m) and a maximal stem diameter of 210 cm (approximate tree height 35 m). We used Enquist's scaling equations to estimate the diameter and height of the crowns of individual trees.

To quantify the relative strength of competition, we estimated the degree of crown overlap between neighboring trees based on their percent crown overlap. *Percent crown overlap* is the area of the ground covered by the overlapping foliage of two individual plants (Figure 1). Because the majority of a plant's root system occurs beneath its crown, percent canopy overlap also reflects the sharing of belowground resources.

To calculate the percent crown overlap of two trees, we took advantage of three observations. The first is the established relationship between stem and crown diameter we just discussed. The second is the known inter-trunk distance of any two trees in the study plot. This distance d calculates as $d = [(x_1 - x_2)^2 + (y_1 - y_2)^2]^{1/2}$, where (x_1, x_2) and (y_1, y_2) are the coordinates of two trees. The third observation regards shape characteristics common to all trees. Specifically, the branches and roots of trees are space-filling systems and trees generally maximize their use of the physical space available to them. To preserve vertical stability and maximize resource use, the crown and root systems of trees tend to grow out from the trunk symmetrically. As a result, canopy cross-sectional areas tend to be circular and can be measured by canopy radius measurements (Niklas 1994; Porter 1989; West et al. 1999). One can thus estimate the area of ground shaded by a tree's crown as the area of a circle with the same diameter as the crown. By extension, one can calculate the percent crown overlap of two trees with the following equation, which describes the area of the lens formed by two intersecting circles:

$$A = r^2 \cos^{-1}\left(\frac{d^2 + r^2 - R^2}{2dr}\right) + R^2 \cos^{-1}\left(\frac{d^2 + R^2 - r^2}{2dR}\right) - \frac{1}{2}\sqrt{(R+r-d)(d+r-R)(d+R-r)(d+r+R)} \quad (1)$$

Here, R is the crown radius of the larger tree, r is the crown radius of the smaller tree, and d is the distance between the crown centers, given by the inter-trunk distance. We calculated percent crown overlap as the percent of the crown area of the smaller tree covered by the crown area of the larger tree. Thus, percent crown overlap is a continuous variable on the interval [0,100].

Species association networks (SANs).

We wanted to ask whether competition manifests itself differently among trees of different body sizes and extent of crown overlap. We thus chose to distinguish trees of different body size classes and tree pairs with different values of percent crown overlap in our representation of tree community structure. We represent the interactions of trees with a given body size and percent crown overlap as a directed graph, a mathematical object that consists of *nodes* and *directed edges*. We call this graph a *species association network*. Briefly, the nodes in this graph are species, and an edge connects species V_1 to species V_2 if the crown of at least one individual of species V_1 overlaps with the crown of at least one individual of species V_2 . More specifically, the construction of a SAN, which we detail below, involves a set of reference trees, which we partition according to stem diameter into four classes: 10-20cm (3,213 individuals, 92 species), 20-30cm (1,163 individuals, 74 species), 30-40cm (577 individuals, 50 species), >40cm (553 individuals, 51 species). Because there are many fewer larger trees than smaller trees, we pooled in the last of these body size classes all individuals with a stem diameter of more than 40cm to give our analysis sufficient statistical power. Similarly, we grouped tree pairs into 10 intervals of percent crown overlap ranging from 0-10 percent crown overlap to 90-100 percent crown overlap (Table 1). In total, we thus generated $4 \times 10 = 40$ SANs, one each for the four different body size and ten different intervals of percent canopy overlap.

We constructed each of these SANs in the following way. We first established a graph that consisted only of isolated nodes – nodes without edges – each of which corresponds to one of the 106 species in the study plot. For each of these nodes (species), we then sequentially carried out the following steps. First, we identified all individuals in the study plot of a given species whose stem diameter fell in the appropriate interval. We call these individuals *reference trees*. Second, for each reference tree, we identified its neighborhood, that is, all trees, regardless of body size, that are not of the same species, and that have a percent crown overlap that falls within the pre-specified interval. For any species V that has a tree in this neighborhood of a reference tree, we establish an edge in the graph from the current focal species (node) to the node corresponding to species V (Figure 1).

The effects of niche differences may increase with tree size and with the extent of shared overlap among trees. However, there are many more small trees than large trees, and many more tree pairs with small than with large crown overlap. These observations raise the question whether we would be able to observe the effects of competition at all, if we had not subdivided the network into trees of different sizes and crown overlap. To address this question, we also generated an “all-neighbors” SAN that included all trees ≥ 5 cm stem size, and all pairwise interactions involving more than 10 percent crown overlap.

We note that some nodes may be isolated in a SAN, corresponding to species where no reference tree has a neighbor within the specified percent shared cover class. This number of isolated nodes can vary according to the tree size and percent crown overlap classes considered.

Measures of Network Topology.

To characterize the topology of our 40 species association networks, we evaluated the following commonly used indicators of graph topology, indicators that are helpful to characterize both biological and other networks (Barabasi et al. 2000; Dunne et al. 2002; Jeong 2000; Strogatz 2001; Watts and Strogatz 1998). We used the LEDA 4.3 package of C++ class libraries (Mehlhorn and Näher 2000) and standard algorithms such as the breadth-first search algorithm (Gibbons 1985) for all our graph-based analyses.

Largest weakly-connected subgraph. A subgraph is a graph that contains a subset of the nodes and edges of a larger graph. An undirected path between two nodes is a sequence of edges that connects the two nodes, regardless of edge directionality. A weakly connected subgraph is a subgraph in which any two nodes (species) are connected via an undirected path. If species are randomly distributed in space, the likelihood that a species is represented in the largest connected subgraph is proportional to its relative abundance in the community. Extreme habitat requirements or strong negative interactions may isolate a species from the rest of the community, reducing the probability that it is represented in the largest subgraph. By contrast, biotic facilitation and shared habitat requirements among species may enhance the number of neighbors a particular species encounters, increasing its chances of being represented in the largest subgraph. Hence differences in subgraph size between random and empirical networks may indicate differences in the probability that species co-occur as neighbors.

Node degree distribution. In a directed graph, a node's degree is the sum of the number of outgoing edges (outdegree d_o) and incoming edges (indegree d_i). The degree distribution characterizes the likelihood that a randomly chosen node will have a degree of a given value. To increase the statistical power of our analysis, we here added the outgoing and incoming edges for each node ($d=d_o+d_i$) and characterized the distribution of d . The node degree represents the number of heterospecifics that each species interacts with in local neighborhoods. Negative interactions (i.e. competition) and strong differences in abiotic preferences among heterospecifics can reduce a species node degree relative to a randomly organized community, while positive interactions (e.g. facilitation) and shared habitat preferences may increase node degree.

Clustering Coefficient (CC). Consider a species V and its k_v neighbors – species directly connected to V – in a SAN. The clustering coefficient of V measures the likelihood that these neighboring nodes or species are also neighbors of each other. Specifically

$$CC_V = \frac{E_V}{k_v(k_v - 1) / 2}$$

where E_V is the number of edges among the k_v species (excluding V itself and regardless of edge directionality). Here, the value $k_v(k_v - 1) / 2$ is the maximum possible number of edges between all k_v species connected to V . The clustering coefficient ranges from zero (i.e. the neighbors of V are not neighbors of each other) to one (i.e. each neighbor of V is also a neighbor of every other neighbor of V). The value of CC_V measures the 'cliquishness' of a graph, that is, the extent to which species form small groups that show preferential interactions within the group. This could occur, for example, if several species were restricted to a specific habitat type, and were thus more likely to be associated in this habitat type. Species that have extreme habitat requirements or that exert strong negative effects on neighbors are more likely to have low values of CC_V . The clustering coefficient of a graph is the average of the clustering coefficients of the individual nodes.

When trees are randomly distributed in space the number of heterospecific neighbors encountered by a species V is proportional to the relative abundance of V . The relative abundance of species in the random SANs we constructed is identical to that of the empirical SANs (see Network Randomization, below). Therefore, higher (lower) empirical values of the average clustering coefficient indicate that certain species are more (less) likely to coexist within local neighborhoods than would be expected in a community of randomly distributed species.

Characteristic Path Length. Path length refers to the shortest distance between two connected nodes, measured by the number of edges that separate them. The characteristic path length, CPL, of a graph is the shortest path length, averaged over all node pairs that are connected by a path of edges. Characteristic path length is a measure of the connectedness of a graph. The more heterospecific neighbors each species in a SAN has, on average, the shorter the characteristic path length of the SAN. Thus factors that enhance the likelihood that species co-occur as neighbors, such as high relative abundance, shared habitat preferences, and facilitation can reduce the path length between any two species (nodes) in the SAN. By contrast, the connectedness of a SAN will be lower, and the characteristic path length will be greater, if competition and abiotic filtering exert a strong negative influence on species co-occurrence.

Network Randomization.

An important aspect of our study was to compare the structure of empirically observed SANs to appropriately randomized networks. To give our analysis statistical power, we generated a set of 1,000 randomized SANs for each of the 40 SANs we had constructed. In order to create any one randomized network, we randomly rearranged the geographic coordinates of all individual trees in the study plot. Specifically, we created a $N \times 4$ matrix where N (the number of rows) is the total number of individual trees and the 4 columns of each row contain the species name, stem diameter, x-y coordinates, and a uniformly distributed random number for each tree (random numbers were generated using the standard ANSI C function `rand()`). We then sorted only the columns containing the x-y coordinates and random numbers by row, in ascending order of random number. The result was a matrix in which each row contains the species name, stem-diameter, and randomly reshuffled x-y coordinates of an individual. We then constructed a SAN from this randomized data in exactly the manner outlined above.

This approach has the advantage of maintaining each species' body size (stem diameter) distribution, which influences the topology of a SAN. At the same time, the approach randomizes the spatial distribution of trees, the aspect of community structure relevant to our study of tree species interactions and community structure. A caveat of this approach is that it may create a biologically unrealistic situation where two very large trees can occur, by chance alone, arbitrarily close to one another. However, because there are relatively few large trees, this scenario is unlikely. For example, fewer than 5% of the trees on the study plot have a stem diameter greater than 40cm.

Regression analysis

To evaluate the effect of percent crown overlap on maximally connected subgraph size, number of edges, characteristic path length, and clustering coefficients of SANs we used a linear regression analysis. In this analysis, which we carried out separately for the four different body size classes, percent crown overlap was the independent variable. We performed an analogous regression analysis for randomized SANs. In this analysis, we used the mean size of maximally connected subgraphs, mean number of edges, mean characteristic path length, and mean

clustering coefficient for 1,000 randomized networks. Where necessary, we logarithmically transformed the data to linearize it prior to regression analysis. If data transformation was not sufficient to remove strong nonlinear behavior, we used polynomial regression to estimate coefficients of determination, R^2 .

Results

As defined in greater detail in Methods, a species association network (SAN) is a graph whose nodes are species. An edge connects species V_1 to species V_2 if the crown of at least one individual of species V_1 overlaps with the crown of at least one individual of species V_2 . Because competition may affect trees of different body size and crown overlap differently, we subdivided the trees in the study plot into several different data sets, corresponding to different SANs. Specifically, we grouped trees into four different body size classes – as indicated by their stem diameter at breast height – and we grouped tree pairs into ten different intervals of percent crown overlap (Table 1). This means that we analyzed $4 \times 10 = 40$ different SANs. Additionally, we constructed a single SAN representing all of the trees on the study plot ≥ 5 cm stem size that had a percent crown overlap of at least 10%.

For each of the 41 SANs, individual species fell into two categories. They were either isolated species, that is, they showed no interactions at all with any other species in the study plot, or they were part of one large subgraph in which any two species could be connected through an undirected path of edges. (A subgraph is a graph containing only some of the nodes and edges of a larger graph.) In other words, all SANs consisted of one connected subgraph involving many species, and several isolated species. Because we were most interested in species interactions, and because only species in the connected subgraph show any interactions, we focus in all analyses below on this connected subgraph and exclude isolated species from the SAN. Below, we first report some descriptive statistics of these SANs. Then we discuss how they differ from randomized SANs, networks in which trees were randomly reshuffled.

Number of species and their interactions in species association networks.

The number of species in the connected subgraph of a SAN was always lower than the 106 total species in the study plot. It decreased with increasing percent crown overlap among species. For 1-10 percent crown overlap, an average (over the four body size classes) of 75.9% of the 106 species were part of the connected subgraph, whereas for 90-100 percent shared cover, on average only 64.9% of the species were part of this subgraph. This downward trend in number of species was statistically highly significant (Table 2). The only systematic exception to it occurred in the 90-100 percent shared cover interval, in which the SANs contained a slightly higher number of species (68.8 species on average) than in the next-lower interval of 80-90 percent shared cover (65.5 species on average). Similarly, the total number of species in a SAN for a given value of percent crown overlap decreased at first for increasing body size, but it increased slightly in the largest body size class. Specifically, the following species numbers – averaged over all intervals of percent crown overlap – were part of the largest connected subgraph: 10-20cm stem diameter: 81.9 species; 20-30cm stem diameter: 72.1 species, 30-40cm: 62.4 species, >40cm: 69.5 species. We currently have no explanation for the slight increase of species in the largest body size class and for species with 90-100 percent shared cover.

The total number of pairwise species associations (edges in a SAN) varied in a similarly predictable way: it generally decreased with increasing body size and with increasing percent

crown overlap. For example, for trees with stem diameter of 10-20cm, we observed 1084 pairwise species associations at 1-10 percent crown overlap, but merely 381 interactions at 90-100 percent crown overlap. For trees with the largest stem diameter (>40cm), we observed 418 pairwise species associations at 1-10 percent crown overlap, but only 221 associations at 90-100 percent crown overlap. The downward trend was again highly statistically significant (Table 2) and consistent across all categories, with only two exceptions. First, SANs with 90-100 crown overlap showed a higher number of associations (249.8, averaged over all body size classes) than the next-lower interval of 80-90 percent shared cover (199.8 associations). Second, SANs in the highest body size class showed more associations (231.6, averaged over all percent crown overlap categories) than the next, 30-40cm stem diameter class (202.4 associations on average). The general downward trend in number of species and species associations may result from the simple fact that there are many fewer large trees than small trees (Figure 2).

Number of species and species interactions are consistent with random community organization.

We next asked whether the number of interacting species and the number of pairwise species associations are significantly different in the 40 empirically observed SANs and in randomized networks, which we obtained by randomly reshuffling tree coordinates in the study plot. If the empirically and randomized communities differ in this regard, community structure may not be explicable by random processes. For each stem diameter class, we first determined the number of species in the ten empirically observed SANs distinguished by their percent crown overlap. We thus obtained a distribution of the number of species over the ten crown overlap categories. We then calculated an analogous distribution for the randomized SANs of the same stem diameter class, where we averaged – again separately for each percent crown overlap category – species numbers over the 1000 randomized SANs. Subsequently, we compared the two distributions via Kolmogorov-Smirnov tests (Sokal and Rohlf 1995) and found no significant differences ($P > 0.3$ for all four stem diameter classes, results not shown). We followed an analogous procedure to compare the number of species associations (edges) in empirically observed and randomized SANs. Again the species associations were not different in the empirically observed and randomized communities ($P > 0.3$ for all four stem diameter classes, results not shown). This indicates that empirically observed SANs have species numbers and numbers of species interactions that are indistinguishable from randomized networks.

The distribution of species associations is consistent with random community organization.

We next assessed community structure through the distribution of the number of species with which any one species interacts. This distribution, the node degree distribution of a SAN, is generally strongly skewed. Specifically, it is dominated by species which interact with few other species (Figure 3). This skewness becomes more pronounced with increasing percent crown overlap (Figure 3), which may result from the fact that there are fewer overall species interactions as percent crown overlap increases (Table 2). We compared the distribution of species interactions between all 40 empirically observed SANs and the corresponding randomized networks. While the two distributions were not absolutely identical – representative examples are shown in Figure 3 – we observed no statistically significant differences in Kolmogorov-Smirnov tests ($P > 0.05$ for each of the 40 SANs). The distribution of the number of species association is thus indistinguishable from a random distribution.

Characteristic path length suggests non-random community organization.

The characteristic path length of a graph is the average shortest path length between all pairs of nodes that are connected through a path of edges, regardless of edge direction. We tested separately for all 40 SANs whether the characteristic path length was significantly different in empirically observed SANs and randomized SANs. To do so, we determined a distribution of characteristic path lengths from the 1000 randomized SANs we had generated for a given empirically observed SAN. We then determined the fraction \mathbf{P} of the values in this distribution that were either greater or less than the empirically observed characteristic path length. This amounts to a two-tailed test of the hypothesis that the empirically observed value is significantly different from the value in randomized networks. The level of confidence in this hypothesis is encapsulated by the value of \mathbf{P} . Figure 4 shows the results of this analysis for the 40 networks representing different body sizes and percent crown overlap. The four panels in the figure are distinguished by tree body size. Error bars plotted on the figures represent 95% confidence intervals of characteristic path lengths. Asterisks above a pair of points indicates that the empirically observed characteristic path length is significantly different from that in randomized SANs. Figure 4 shows multiple instances of significantly different path lengths, which we take to be an indication of non-random community organization.

For all body size classes and both in empirically observed and randomized SANs, there is a statistically significant increase in characteristic path length with increased crown overlap (percent crown overlap; Figure 4; Table 3). This increase is readily explained, if one considers that a graph's characteristic path length generally increases as its total number of edges (species interactions) decreases. Such a decrease of total species interactions is precisely what occurs as crown overlap increases (Table 2). We asked whether the distribution of characteristic path length across different crown overlap categories is significantly different for empirically observed and randomized SANs. The answer is no (Kolmogorov-Smirnov test; $\mathbf{P}>0.1$ for all body size classes; Figure 3; Table 3). Thus, while characteristic path length indicates non-random community organization for some intervals of percent shared cover, such non-random organization is not evident across the whole range of percent crown overlap.

Clustering coefficients strongly suggests non-random community organization.

Exactly analogous to our analysis of characteristic path lengths, we tested separately all 40 SANs whether the clustering coefficient was significantly different in empirically observed SANs and randomized SANs. The clustering coefficient (defined in the Methods) indicates the preferential association of species within a small species groups. The results of this analysis are shown in Figure 4, whose format is analogous to Figure 3. The analysis shows that for most intervals of crown overlap and tree size, clustering coefficients are significantly higher in empirically observed SANs than in random SANs, at \mathbf{P} -values that often exceed 10^{-4} . Specifically 35 of 40 SANs showed significantly higher clustering coefficients at $\mathbf{P}<0.01$, and 28 of 40 networks showed higher clustering coefficients at $\mathbf{P}<10^{-4}$. These differences become more pronounced as body size and crown overlap increases (Figure 5). That is, clustering coefficients tend to be dramatically higher for large trees with a high percent crown overlap.

For all body size classes and both in empirically observed and randomized SANs, we see a statistically significant decrease in clustering coefficient with increased crown overlap (percent crown overlap; Figure 5; Table 3). This increase is again readily explained, if one considers that clustering coefficients tend to decrease as a network becomes sparser, that is, as the number of edges (species interactions) decreases. Such a decrease of total species interactions

is precisely what occurs as crown overlap increases (Table 2). We also asked whether the distribution of clustering coefficients across different crown overlap categories is significantly different for empirically observed and randomized SANs. As opposed to the analogous analysis for characteristic path length (Figure 4), we observe significant differences for all but the smallest tree size categories (Figure 5; Table 3).

Effects of competition are visible in SANs involving all trees and interactions.

We asked whether the various indicators of network topology we analyzed above were significantly different from their counterparts in randomized pan-neighborhood networks. The number of species in the empirically observed SAN (105) was statistically indistinguishable to its average value of the randomized networks of 104.68 ($P=0.189$). The same was true for the total number of species associations. We observed 3508 species associations in the empirically observed pan-neighborhood, compared to 3769.44 species associations in the randomized pan-neighborhood SAN ($P=0.999$). The distributions of the number of interactions per species were not significantly different in the empirically observed and in the randomized pan-neighborhood SANs (Kolmogorov-Smirnov test; $P=0.9655$). In contrast, the characteristic path length in the empirically observed pan-neighborhood SAN was significantly higher than in the randomized SAN (empirical: 1.78 random: 1.69; $P<10^{-4}$), and the clustering coefficient was significantly lower in the empirically observed pan-neighborhood SAN (empirical =0.7433; random =0.7990; $P<10^{-4}$).

Overall, we concluded from this analysis that competition manifests itself no less strongly in the pan-neighborhood network than in many of the 40 SANs distinguished by tree body size and crown overlap.

Discussion

We evaluated the effect of interspecific competition on network structure by constructing species networks on the basis of the crown overlap of neighboring trees. While the non-random structure of the species networks we observe is consistent with the effects of competition, the spatial distribution of species may also have been influenced by abiotic factors and species interactions other than competition. For example, edaphic conditions vary across the study site, with some parts being drier than others. The specific environmental conditions needed for germination and establishment differ among tree species (Clark and Clark 1992), and those with narrow habitat requirements could potentially have been constrained to specific areas of the site that support those requirements. With regard to non-competitive species interactions, the establishment of some individuals may actually be facilitated by the presence of other species (Connell and Slatyer 1977; Langenheim et al. 1984). In addition, predators attracted to a given species may also reduce the local density of neighboring heterospecifics (Connell 1971; Janzen 1970), inducing spatial correlation among some species. Our goal was not to establish the preeminence of competition as the structuring force, but to determine whether community structure was influenced by species niche differences as opposed to neutral processes of random birth, death, and dispersal.

The use of graph-theoretic tools has precedents in several areas of biology. Prominent examples include cell biology (Fell and Wagner 2000; Jeong 2000; Wagner 2001) and ecology (Cohen and Briand 1984; Cohen and Palka 1990; Dunne et al. 2002; Montoya and Sole 2002; Pimm 1991; Williams et al. 2002). Most ecological examples of graph-based approaches come

from food web analysis, which focuses on the properties of networks that represent the trophic interactions among species. However, graph-based analyses also have great potential for analyzing communities that represent a single trophic group, such as our study community. There are three merits of graph-based approaches in this area. The first is that a graph representation is finer-grained than traditional representations of community structure, such as species abundance distributions. Second, graph representations are highly flexible and can represent community structure at many different levels of resolution. We here used a low-level representation, a species association network whose nodes are species – not individuals. (The appropriate choice of representation may be critical to detect the effects of competition on communities.) A third and related motivation for graph-based analyses is that any one graph can be characterized in many ways, and thus many ways in which it could reveal the effects of competition on a community. Again, the choice of characterization method is crucial here. We used only a few different ways to characterize the structure of SANs, some of which revealed nonrandom community structures, whereas others did not. There may be many other, more sensitive indicators of competition in communities, indicators that have not yet been discovered. Our study is only the first step into a rich new field of investigation.

Some observations that emerge from our analysis are unsurprising and can be explained by simple observations about community structure. One such observation is that the number of species in a SAN and the number of pairwise associations between them is negatively correlated with the body size of trees. In steady-state forests, the number of individual trees in a sample area is inversely proportional to body size (Enquist and Niklas 2001; Niklas 1994). The number of tree species in the sample area is directly proportional to the number of individuals (Fisher et al. 1943; MacArthur and Wilson 1967; Preston 1948; Rosenzweig 1995). Thus, the large trees in a sample area like our study plot will contain representatives of fewer species than the smaller trees. This will naturally lead to fewer species and interactions per SAN.

Another readily explained observation is that the number of species in a SAN and the number of pairwise associations between them decreases as the percent crown overlap between tree pairs increases. This correlation is again a consequence of the relation between tree number and species richness in a sample area. A group of trees with a given size but large crown overlap will occupy a smaller area than a group of trees with small crown overlap. Because a smaller area will contain fewer species than a large area, the total number of interacting species and interactions will be reduced for trees with large crown overlap.

A third readily explained observation is the dependence of characteristic path length and clustering coefficient on percent crown overlap. In most graphs of a given structure, a thinning of the graph through edge removal will lead to an increase in the graph's characteristic path length and a decrease in the clustering coefficient. Such a thinning of edges effectively occurs as the fraction of percent crown overlap increases, simply because there are fewer trees with large crown overlaps.

The essence of our approach to detect potential effects of competition is to compare the structure of an empirically observed tree community with that of a randomized community in which the locations of individual trees are randomly reshuffled. Some indicators of community structure such as the distribution of the number of interactions per species revealed no differences between empirically observed and randomized communities. However, two indicators showed such differences. First, the characteristic path length of some SANs is higher than that of randomized communities, a difference that tends to increase with increasing crown overlap and that is statistically significant for 18 out of 40 SANs. Second, the clustering

coefficient of most SANs is significantly higher than that of randomized SANs. Its difference to random networks increases with increasing tree size and with increasing percent crown overlap. The clustering coefficient, which measures the preferential association of groups of species, is thus the most effective indicator of nonrandom community organization for the community we analyzed.

Taken together, the higher empirical values of characteristic path length and clustering coefficient suggest that crown overlap and tree size influence the ability of the tree species to coexist as neighbors. The elevated values of characteristic path length of the empirical SANs suggest that some species are less likely to be represented in any given local neighborhood, on average, than expected if they were equal competitors. Yet the higher clustering coefficients suggest that certain species are *more* likely to co-occur *with each other*. There are many possible biological explanations for these patterns which can be broadly attributed to niche differences and niche similarities among species.

Non-uniform spatial distributions of individuals in one species can arise through competition (Clark and Clark 1992), but also through random processes. For example, Hubbell suggested that clumped distributions of trees could arise if mature trees act as point sources for the dispersal of seeds that germinate to produce offspring (Hubbell 1979). In this scenario, the likelihood of encountering a tree of the same species declines with increasing distance from a parent tree, a pattern of distribution known as spatial autocorrelation. This explanation relies only on the random dispersal of seeds and fruits. By extension, could the heterospecific associations we see be caused by similar random processes? We believe on two counts that this is highly unlikely. First, indicators of network topology such as the clustering coefficient are much higher in empirically observed communities than in randomized communities, which is not consistent with this possibility. It is also highly unlikely that processes of random seed dispersal within a species can cause the highly significant heterospecific associations we see. Such processes, which occur independently among species, would not lead to significant cross-correlations in the spatial distribution among species. Finally and most importantly, non-randomness in community structure increases with increasing tree size and thus age. If a process such as random seed dispersal were to cause the associations we see, then this process should show its effects across all tree sizes. Yet this is not what we observe. In contrast, increased non-randomness of community structure with increasing tree size is perfectly consistent with a competition-based explanation of community structure. The reason is that older trees contain a longer historical record of competition than younger trees, which suggests that effects of competition on a community may be cumulative in time, as we observe.

In closing, we note an intriguing parallel of our tree community with small world graphs, types of networks observed in many natural systems from cells to human societies. Small world networks are characterized by a high clustering coefficient and a small characteristic path length. Most direct interactions in such networks occur within groups of individuals that are highly connected. Nonetheless, in such networks one can reach any node from any other node by a very short path of edges (Watts 1999; Watts and Strogatz 1998). Our tree community is such a small-world network. In this network, the biological significance of the clustering coefficient is clear: It represents preferential associations of species groups. The biological significance of the short characteristic path length is much less clear. We have used it merely as a statistical indicator of non-random community structure. However, we speculate on a possible consequence of such a short characteristic path length. A short characteristic path length causes information to spread very rapidly through a graph. In tropical forests, the rapid transfer of information regarding plant

phenology among species could facilitate the ability of species to synchronize the timing of flowering or fruiting. This is just another example of how a graph-representation of biological communities may lead to new perspectives on biological phenomena.

Acknowledgements

This research was supported by a NSF Research Training Grant (DGE 9553623) and Biocomplexity grant (DEB 0083422) to James H. Brown and Bruce Milne. We are indebted to James H. Brown for support and guidance during all stages of this research. MF is indebted to Susan Fuller for her support. AW would like to thank the Santa Fe Institute for its continued support. We are also sincerely grateful to Gavin Conant, Mike Gilchrist, and Annette Evangelisti for their input.

References Cited

- Barabasi, A. L., R. Albert, and H. Jeong. 2000. Scale-free characteristics of random networks: the topology of the World-Wide Web. *Physica a* ; 281:69-77.
- Bell, G. 2000. The distribution of abundance in neutral communities. *The American Naturalist* 155:606-617.
- . 2001. Neutral Macroecology. *Science* 293:2413-2418.
- Brown, J. H. 1984. On the relationship between abundance and distribution fo species. *The American Naturalist* 124:255-279.
- Cantwell, M. D., and R. T. T. Forman. 1993. Landscape graphs: Ecological modeling with graph theory to detect configurations common to diverse landscapes. *Landscape Ecology* 8:239-255.
- Clark, D. A., and D. B. Clark. 1992. Life-History Diversity of Canopy and Emergent Trees in a Neotropical Rain-Forest. *Ecological Monographs* ; 62:315-344.
- Cohen, J. E., and F. Briand. 1984. Trophic linnks of community food webs. *Proceedings of the National Academy of Sciences USA* 81:4105-4109.
- Cohen, J. E., and C. M. Newman. 1985. A stochastic theory of community food webs: I. models and aggregated data. *Proceedings of the Royal Society of London Series B Biological Sciences* 224:421-448.
- Cohen, J. E., and Z. J. Palka. 1990. A Stochastic Theory of Community Food Webs. V. Intervality and Triangulation in the Trophic-Niche Overlap Graph. *American Naturalist* ; 135:435-463.
- Connell, J. H. 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees., Pages 298-312 *in* P. J. den Boer, and G. R. Gradwell, eds. *Dynamics of numbers in populations*. *Proceedings of the Advanced Study Institute on dynamics in numbers in populations*, Oosterbeek, 1970. Wageningen, Centre for Agricultural Publising and Documentation.
- Connell, J. H., and R. O. Slatyer. 1977. Mechanisms of succession in natural communities and their role in community stability and organization. *The American Naturalist* 111:1119-1144.
- Duivenvoorden, J. F., J. C. Svenning, and S. J. Wright. 2002. Ecology : Beta diversity in tropical forests. *Science* ; 295:636-637.
- Dunne, J. A., R. J. Williams, and N. D. Martinez. 2002. Food-web structure and network theory: the role of connectance and size. *Proceedings of the National Academy of Sciences USA* 99:12917-12922.
- Enquist, B. J., and K. J. Niklas. 2001. Invariant scaling relations across tree-dominated communities. *Nature* 410:655-660.
- Enquist, B. J., G. B. West, E. L. Charnov, and J. H. Brown. 1999. Allometric scaling of production and life-history variation in vascular plants. *Nature* 401:907-911.
- Fell, D., and A. Wagner. 2000. The small world of metabolism. *Nature Biotechnology* 18:1121-1122.
- Fisher, R. A., A. S. Corbet, and C. B. Williams. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Ecology* 12:42-58.
- Gibbons, A. 1985, *Algorithmic Graph Theory*. Cambridge, UK, Cambridge University Press.
- Harper, J. L. 1977, *Population Biology of Plants*. London, Academic Press.

- Hubbell, S. 1979. Tree dispersion, abundance, and diversity in a tropical dry forest. *Science* 203:1299-1309.
- Hubbell, S. P. 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton, Princeton University Press.
- Hubbell, S. P., and R. B. Foster. 1986. Biology, chance, and history and the structure of tropical rain forest tree communities, Pages 665 *in* J. Diamond, and T. J. Case, eds. *Community Ecology*. New York, Harper and Row.
- Janzen, D. H. 1970. Herbivores and the number of tree species in tropical forests. *American Naturalist* ; 104:501-528.
- . 1986, *Guanacaste National Park: Tropical Ecological and Cultural Restoration*. San Jose, Costa Rica, Editorial Universidad Estatal a Distancia.
- Jeong, H., Tombor, B. Albert, R. Oltvai, Z.N., Barabasi, A.L. 2000. The large-scale organization of metabolic networks. *Nature*. 407:651-654.
- Langenheim, J. H., C. B. Osmond, A. Brooks, and P. J. Ferrar. 1984. Photosynthetic responses to light in seedlings of selected Amazonian and Australian rainforest tree species. *Oecologia* 63:215-224.
- MacArthur, R. H., and E. O. Wilson. 1967, *The Theory of Island Biogeography: Monographs in Population Biology*, v. 1. Princeton, Princeton University Press.
- Mehlhorn, K., and S. Näher. 2000, *LEDA A platform for combinatorial and geometric computing*. Cambridge, England, Cambridge University Press.
- Montoya, J. M., and R. V. Sole. 2002. Small world patterns in food webs. *Journal of Theoretical Biology* 214:405-412.
- Niklas, K. 1994, *Plant Allometry*. Chicago, University of Chicago Press.
- Pimm, S. L. 1991, *Balance of Nature? Chicago*, University of Chicago Press.
- Porter, J. R. 1989. Modules, models and meristems in plant architecture, Pages 143-159 *in* G. Russell, B. Marshall, and P. G. Jarvis, eds. *Plant Canopies: Their Growth, Form and Function*. Society for Experimental Biology, Seminar Series. Cambridge, Cambridge University Press.
- Preston, F. W. 1948. The commonness, and rarity, of species. *Ecology* 29:254-283.
- Rosenzweig, M. L. 1995, *Species Diversity in Space and Time*. Cambridge, UK, Cambridge University Press.
- Sokal, R. R., and F. J. Rohlf. 1995, *Biometry*. New York, W.H. Freeman and Company.
- Strogatz, S. H. 2001. Exploring complex networks. *Nature* 410:268-276.
- Tilman, D. 2000. Causes, consequences and ethics of biodiversity. *Nature* 405:208-211.
- Urban, D., and T. Keitt. 2001. Landscape connectivity: a graph-theoretic perspective. *Ecology* 82:1205-1218.
- Volkov, I., J. R. Banavar, S. P. Hubbell, and A. Maritan. 2003. Neutral theory and relative species abundance in ecology. *Nature* 424:1035-1037.
- Volkov, I., J. R. Banavar, A. Maritan, and S. P. Hubbell. 2004. The stability of forest biodiversity. *Nature* 427:696-697.
- Wagner, A. 2001. How to reconstruct a genetic network from n single-gene perturbations in fewer than n^2 easy steps. *Bioinformatics* 17:1183-1197.
- Watts, D. J. 1999. Networks, dynamics, and small-world phenomenon. *American Journal of Sociology* 105:493-527.
- Watts, D. J., and S. H. Strogatz. 1998. Collective dynamics of 'small-world' networks. *Nature* 393:440-442.

- West, G., J. Brown, and B. Enquist. 1999. A general model for the structure of plant vascular systems. *Nature* 400:664-667.
- Williams, R. J., E. L. Berlow, J. A. Dunne, A. L. Barabasi, and N. D. Martinez. 2002. Two degrees of separation in complex food webs. *Proceedings of the National Academy of Sciences USA* 99:12913-12916.

TABLE CAPTIONS

Table 1. Stem Size Classes and Percent Crown Overlap Levels

To construct species networks, we used trees representing the 4 stem size classes (measured as diameter at breast height) and 10 intervals of shared crown overlap (percent of the crown of the smaller tree overlapped by the crown of the larger tree) shown in this table. We constructed 40 separate species association networks, one for each combination of stem size and crown overlap.

Table 2. Effect of Percent Crown Overlap on Species Number and Pairwise Associations

Shown are the results of linear regression analyses of the effect of shared crown overlap on the number of species and number of pairwise species associations represented in the weakly-connected subgraph of each species association network (SAN). The values of regression R^2 , F, and P are shown for separate regression analyses of the empirically observed and randomized SANs in each of the 4 stem size classes. To calculate values of R^2 , F, and P for randomized SANs, we used the average number of species and average number of pairwise associations over 1000 random SANs as dependent variables. The column entitled “KS Emp vs Ran” contains the results of Kolmogorov-Smirnov tests of the distributions (not shown) of empirical and average randomized species numbers and numbers of pairwise associations over 10 categories of percent crown overlap.

Table 3. Effect of Percent Crown Overlap on Characteristic Path Length and Clustering Coefficient

Shown are the results of linear regression analyses of the effect of shared crown overlap on the characteristic path length and clustering coefficients of each species association network (SAN). The values of regression R^2 , F, and P are shown for separate regression analyses of the empirically observed and randomized SANs in each of the 4 stem size classes. To calculate values of R^2 , F, and P for randomized SANs, we used the average characteristic path length and clustering coefficients over 1000 random SANs as dependent variables. The column entitled “KS Emp vs Ran” contains the results of Kolmogorov-Smirnov tests of the distributions (not shown) of empirical and average randomized characteristic path length and clustering coefficients over 10 categories of percent crown overlap.

FIGURE CAPTIONS

Figure 1. Constructing a Species Association Network

This figure uses a hypothetical example to show how we constructed a species association network where the reference trees (see Methods) had a stem diameter of at least 40cm and where the required percent crown overlap was 20-30%. **1a**: The neighborhoods of two large trees (stem diameter >40cm) that represent two species (labeled in bold-italics **A** and **E**) are shown. All trees that overlap the crowns of **A** and **E** by 20-30% are considered their neighbors (shown with dark shading and labeled with species names: B, C, D, E, F, G, H). The other (non-shaded) trees are not considered for this particular network because their crowns overlap with **A** and **E** by less than 20% or by more than 30%. For example, the small tree just to the left of **A** has a crown overlap with **A** of 100%. **1b**: Aerial view of 1a, showing how the crowns of neighboring trees overlap the crowns of **A** and **E**. Each crown shown corresponds to a tree in figure 1a. Note that the tree belonging to species D falls within the neighborhoods of both **A** and **E**. **1c**: The species network based on 1a and 1b. Nodes represent species. We draw an edge from **A** (**E**) to a species if the species has at least one tree whose crown overlaps with **A** (**E**) by 20-30 percent.

Figure 2. Effect of Stem Diameter on Number of Individual Trees and Number of Species

This figure shows how increasing stem size affects counts of the number of individual trees and species on the study plot. Shaded bars = counts of individuals (left axis); open bars = counts of species (right axis). The number of individuals and species represented in each stem size class decreased as a power function of stem size. Note that the last stem-size classes represents the pooled values of stem sizes from 100-200cm.

Figures 3. Selected Distributions of the Number of Interactions per Species

The distribution of the number of interactions per species varied among the different species associations networks we analyzed. Shown are data for four networks that are representative of the range of distributions observed. In each of the figures, the distributions of empirically observed species association networks (light bars) are compared to those of randomized species association networks (dark bars). The stem size and crown overlap of the networks are as follows: **a** = 10-20 cm, 1-10% overlap; **b** = 10-20 cm, 91-100% overlap; **c** = 40+ cm, 1-10% overlap; **d** = 40+ cm, 91-100% overlap. Note that the distributions become more right-skewed with increasing crown overlap. The larger stem size classes contain fewer species overall.

Figure 4. Effect of Stem Size & Percent Crown Overlap on Characteristic Path Length

Each panel in this figure shows the effect of shared crown overlap (horizontal axis) on the characteristic path length (vertical axis) between species in a network for empirically observed and randomized species association networks (SANs). Each panel represents a different stem size class of SAN. Solid circles show empirically observed values; open squares show average of values in 1000 randomized networks; solid lines indicate the regression fit for empirically observed networks, dashed lines indicate the regression fit for randomized networks; error bars indicate 95% confidence limits. Statistically significant differences between empirically observed and randomized values are indicated by asterisks. No asterisk: $P > 0.05$; '*': $P < 0.05$; '**': $P < 0.01$; '***': $P < 0.001$. Also shown are the P -values of Kolmogorov-Smirnov tests (KS P) that tested whether the distribution of empirical data points of a given stem size class was the same as those of the randomized SAN of the same class. See text for description of results.

Figure 5. Effect of Stem Size & Percent Crown Overlap on Clustering Coefficient

Each panel in this figure shows the effect of shared crown overlap (horizontal axis) on the base-10 logarithm of the clustering coefficient (vertical axis) for empirically observed and randomized species association networks (SANs). Each panel represents a different stem size class of SAN, as indicated in the panels. Solid circles show empirically observed values; open squares show average of values in 1000 randomized networks; solid lines indicate the regression fit for empirically observed networks, dashed lines indicate the regression fit for randomized networks; error bars indicate 95% confidence limits. Statistically significant differences between empirically observed and randomized values are indicated by asterisks. No asterisk: $P > 0.05$; '*': $P < 0.05$; '**': $P < 0.01$; '***': $P < 0.001$. Also shown are the P -values of Kolmogorov-Smirnov tests (KS P) that tested whether the distribution of empirical data points of a given stem size class was the same as those of the randomized SAN of the same class. See text for description of results.

1
2
3
4

Table 1. Stem Size Classes and Shared Crown Overlap Levels

Stem Size (cm)	Shared Crown Overlap Interval (%)									
10 – 20	1 – 10	11 – 20	21 – 30	31 – 40	41 – 50	51 – 60	61 – 70	71 – 80	81 – 90	91 – 100
20 - 30	1 – 10	11 – 20	21 – 30	31 – 40	41 – 50	51 – 60	61 – 70	71 – 80	81 – 90	91 – 100
30 - 40	1 – 10	11 – 20	21 – 30	31 – 40	41 – 50	51 – 60	61 – 70	71 – 80	81 – 90	91 – 100
≥40	1 – 10	11 – 20	21 – 30	31 – 40	41 – 50	51 – 60	61 – 70	71 – 80	81 – 90	91 – 100

5
6

1 **Table 2. Effect of Percent Crown Overlap on Species Number and Pairwise Associations**

2

Stem Size (cm)	SAN	Number of Species				Number of Pairwise Associations			
		R ²	F	P	KS Emp vs Ran	R ²	F	P	KS Emp vs Ran
10 to 20	Emp	0.91	82.42	0.0000	P=0.6751	0.80	31.79	0.0005	P=0.6751
	Ran	0.92	90.73	0.0000		0.81	34.10	0.0004	
20 to 30	Emp	0.73	21.98	0.0016	P=0.3129	0.71	19.95	0.0021	P=0.6751
	Ran	0.81	33.80	0.0004		0.76	25.80	0.0010	
30 to 40	Emp	0.44	6.34	0.0360	P=0.3129	0.67	16.22	0.0038	P=0.6751
	Ran	0.72	20.08	0.0021		0.71	19.36	0.0023	
≥40	Emp	0.51	8.27	0.0206	P=0.9748	0.59	11.63	0.0092	P=0.3130
	Ran	0.49	7.75	0.0238		0.55	9.65	0.0145	

3

1 **Table 3. Effect of Percent Crown Overlap on Characteristic Path Length and Clustering Coefficient**

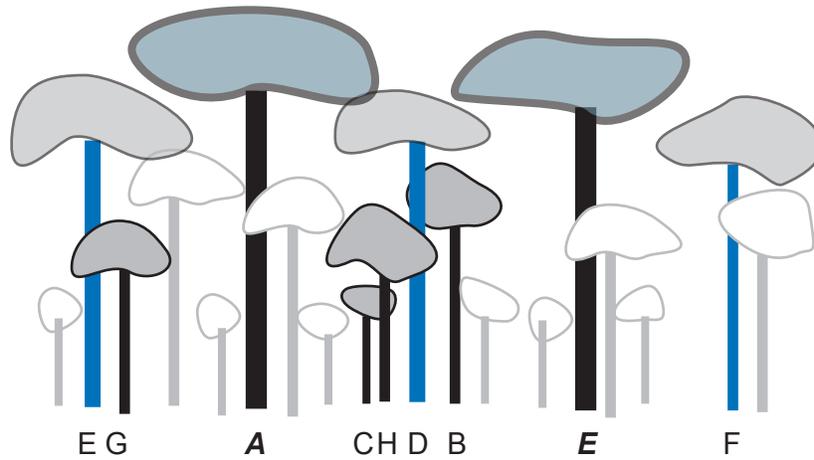
2

Stem Size (cm)	SAN	Characteristic Path Length				Clustering Coefficient			
		R ²	F	P	KS Emp vs Ran	R ²	F	P	KS Emp vs Ran
10-20	Emp	0.81	34.32	0.0004	P=0.3129	0.72	20.63	0.0019	P=0.1108
	Ran	0.92	91.04	0.0000		0.93	99.38	0.0000	
20-30	Emp	0.69	17.61	0.0030	P=0.6751	0.80	32.07	0.0005	P=0.0069
	Ran	0.94	116.98	0.0000		0.90	70.82	0.0000	
30-40	Emp	0.67	15.98	0.0040	P=0.9748	0.65	14.90	0.0048	P<0.0002
	Ran	0.93	105.47	0.0000		0.88	59.56	0.0001	
≥40	Emp	0.47	7.12	0.0284	P=0.1108	0.45	6.52	0.0340	P<0.0002
	Ran	0.75	23.78	0.0012		0.72	21.05	0.0018	

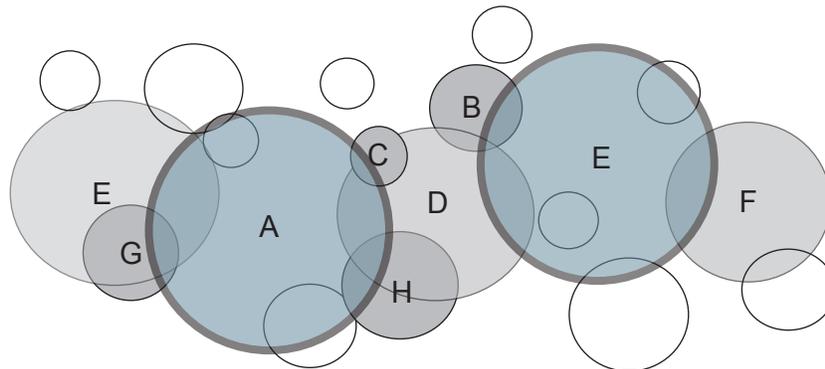
3

Figure 1. Constructing a Species Association Network

1a Local region of forest that includes two large trees (A and E).



1b Aerial view of trees in 1a.



1c Species network constructed from trees in 1a.

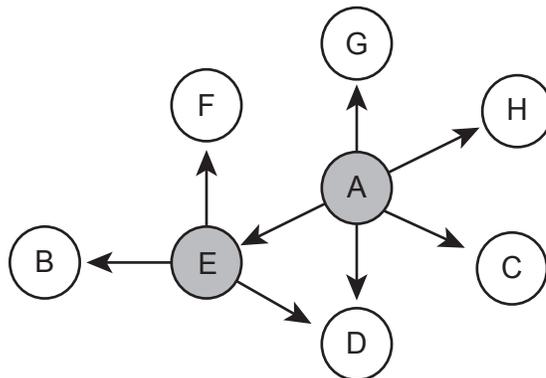


Figure 2. Effect of Stem Diameter on Number of Individual Trees and Number of Species

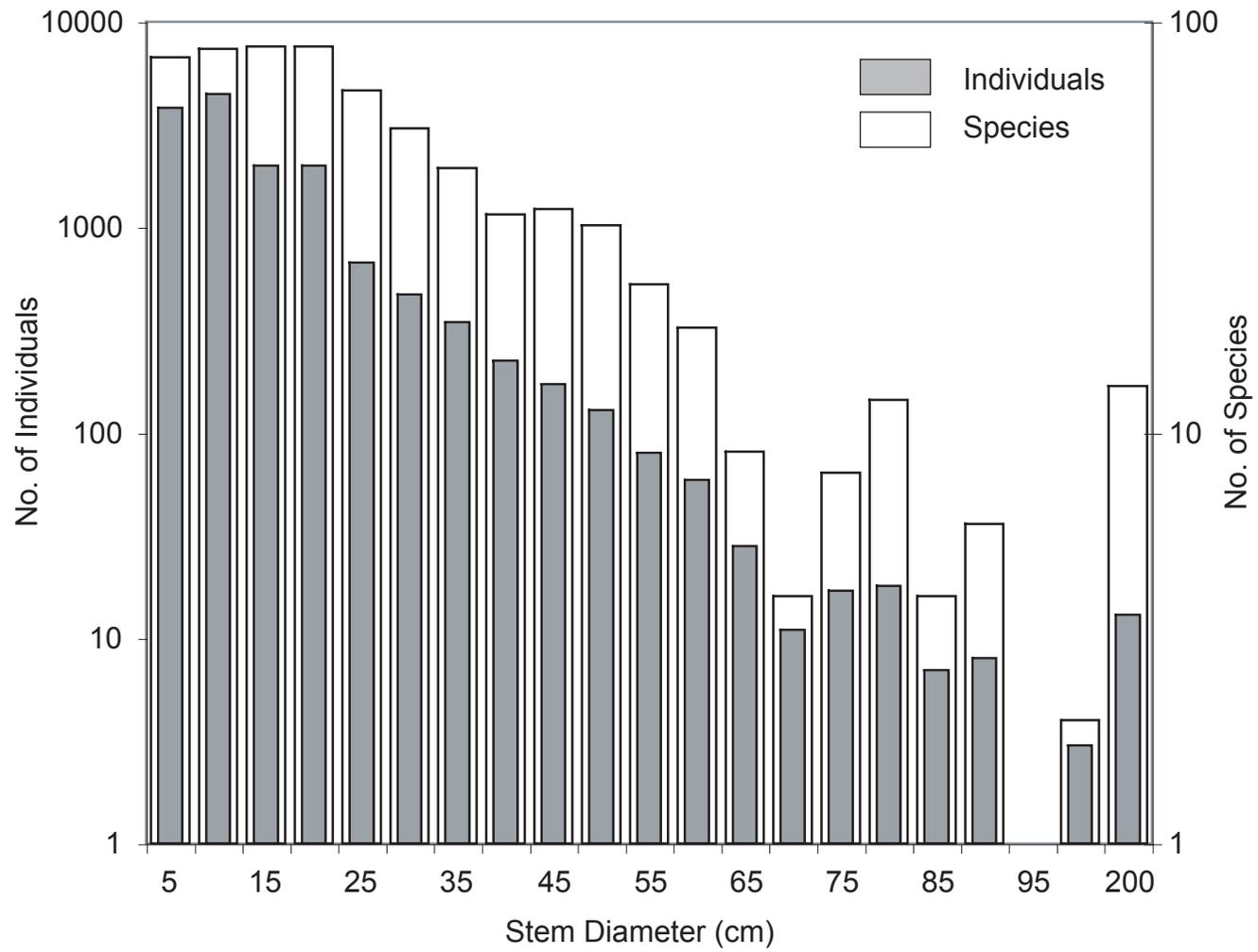


Figure 4. Effect of Stem Size & Shared Crown Overlap on Characteristic Path Length

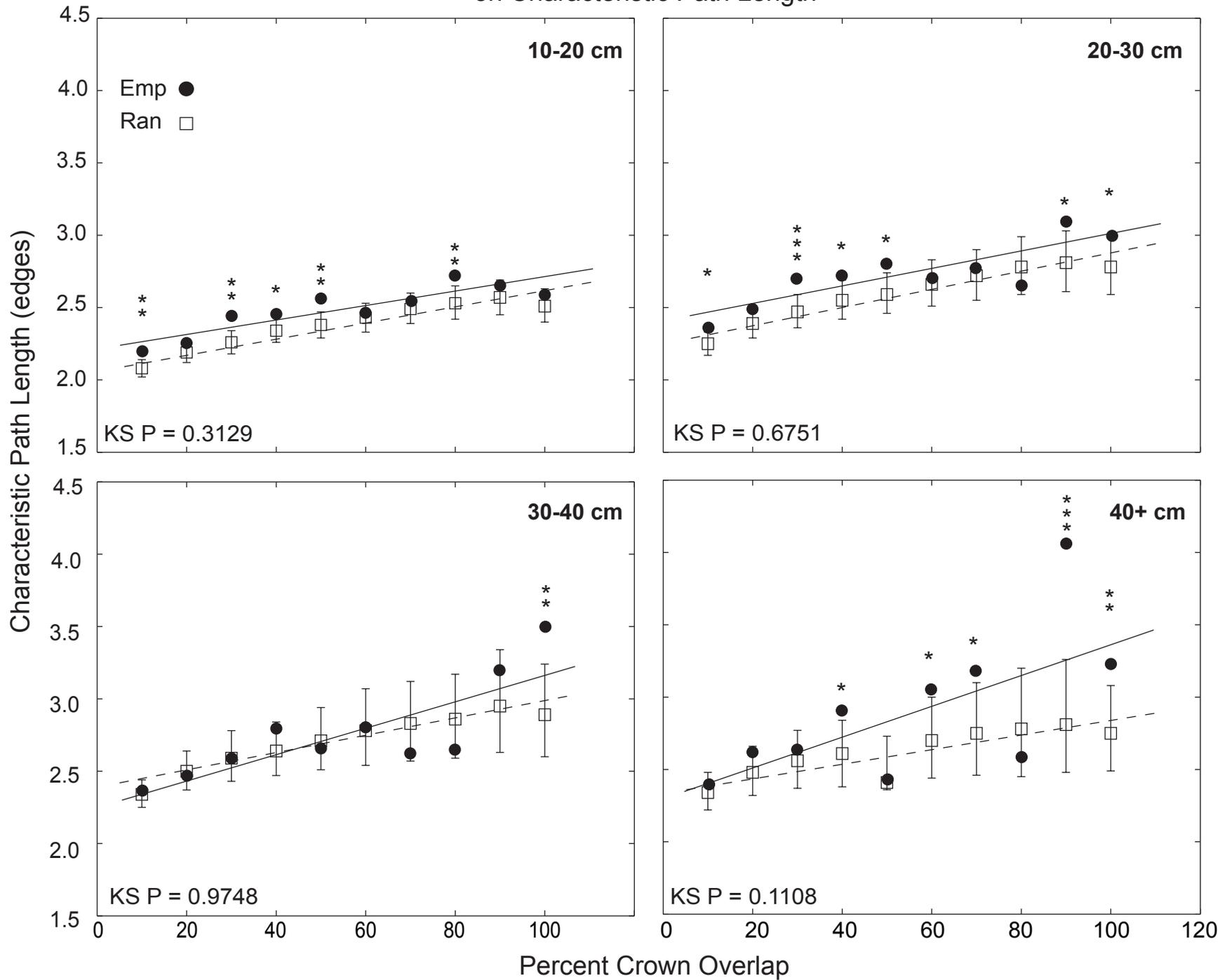


Figure 5. Effect of Stem Size & Percent Crown Overlap on Clustering Coefficient

