The relative contribution of abundance and phylogeny to the structure of plant facilitation networks

Miguel Verdú and Alfonso Valiente-Banuet

The structure of the real ecological networks is determined by multiple factors including neutral processes, the relative abundances of species, and the phylogenetic relationships of the interacting species. Previous efforts directed to analyze the relative contribution of these factors to network structure have not been able to fully incorporate the phylogenetic relationships between the interacting species. This limitation stems from the difficulty of predicting interaction probabilities based on the independent phylogenies of interacting species (e.g. plants and animals). This is not the case for plant facilitation networks, where nurse and facilitated species evolve in a common phylogeny (e.g. spermatophyte phylogeny). Facilitation networks are characterized by both high nestedness and interactions tending to occur between distantly related nurse and facilitated species. We evaluate the relative contribution of phylogeny and species abundance to explain both the frequency of observed interactions as well as the network structure in a real plant facilitation network at Tehuacán Valley (central Mexico). Our results show that the combined effects of phylogeny and species abundance were, by far, the best predictors of both the frequency of the interactions observed in this community and the parameters (nestedness and connectance) defining the network structure. This finding indicates that species interact proportionally to both their phylogenetic distances and abundances simultaneously. In short, the phylogenetic history of species, acting together with other ecological factors, has a pervasive influence in the structure of ecological networks.

Community ecology is experiencing a renewal of interest in the study of the structure of ecological communities thanks to the development of complex network theory (Bascompte and Jordano 2007) and phylogenetic methods (Webb et al. 2002). The few attempts to integrate both fields unambiguously show that phylogenetic effects are important determinants of community structure (Cattin et al. 2004, Rezende et al. 2007, 2009, Bersier and Kehrli 2008). However all these empirical studies also recognize that the variability seen in the interactions between species in an ecological network cannot be explained fully by phylogenetic effects.

Other effects like species abundance or phenotypic complementarity can theoretically explain the structure of the ecological networks, specially the nested organization of the interaction matrix. Differences in species abundance may produce high nestedness (i.e. a tendency of specialists to interact preferentially with generalists) because the most abundant species simply have higher chances to interact with their more abundant counterparts (Vázquez 2005). Similarly, phenotypic complementarity determines the degree of functional matching between interacting species and may thus generate nestedness (Rezende et al. 2007). Interestingly, when traits determining such phenotypic complementarity evolve in a conserved way, the hierarchical nature of phylogenetic relations can contribute to increased nestedness.

Plant facilitation networks are characterized by both high nestedness (Verdú and Valiente-Banuet 2008) and a clear phylogenetic relationship between benefactor (nurses) and beneficiary (facilitated) species (Valiente-Banuet and Verdú 2007, 2008). Plant facilitation interactions are phylogenetically structured because they tend to occur between distantly related nurses and facilitated species. This assembly process results from the fact that species with similar regeneration niches tend to compete whereas species with complementary niches tend to coexist. As the regeneration niche is an evolutionarily conserved trait, coexistence mediated by facilitation is more likely between distantly than closely related species, ultimately leading to phylogenetic overdispersion (Valiente-Banuet and Verdú 2007, 2008). Renegeration niche is ultimately determined by a complex array of phenotypic characters and therefore, niche complementarity is produced by complementarity in many phenotypic traits. Following Rezende et al. (2007) rationale on phenotypic complementarity and phylogenetic conservatism, it would be expected that the high level of nestedness detected in facilitation networks would be explained by the hierarchical structure of phylogenetic relationships derived from the
niche complementarity existing between nurses and their facilitated plant species.

Despite the undeniable value of these theoretical studies, they have not addressed the multiplicity of factors acting simultaneously in the structure of real networks. A recent method developed by Vázquez et al. (2009) provides us with the opportunity to evaluate the relative contribution of multiple factors to explain the observed matrix of interactions as well as the structure of the interaction network (i.e. network connectance – or the proportion of realized interspecific links-, nestedness, etc). The method first derives interaction probability matrices resulting from different factors (i.e. abundance, phenological synchrony, etc) and then calculates the likelihood of these matrices given the observed matrix of interactions. As expected, Vázquez et al. (2009) found that the combination of different effects (abundance and spatiotemporal overlap) was the best predictor of the observed pairwise interactions between pollinators and plants. Despite previous findings that phylogenetic relationships are crucial in shaping community structure, Vázquez et al. (2009) could not incorporate phylogenetic information into their framework because there is no way of deriving an expected probability matrix based on the independent phylogenies of plants and animals. However, the situation is different for ecological interactions between species evolved in a common phylogeny. This is the case for plant facilitation networks because the interactions occur between plants (nurses and facilitated species). Thus, we could easily derive the probability matrix of interactions from the matrix of phylogenetic distances between the nurse and the facilitated species under the assumption of phylogenetic overdispersion (Fig. 1). As facilitation interactions tend to occur between distantly related species (Valiente-Banuet and Verdú 2007, Verdú et al. 2009), we have an exceptional opportunity to evaluate the relative contribution of phylogeny together with other factors, such as abundance, in the pattern of observed interactions. We predict that, together with abundance, phylogenetic relationships between species are important determinants of the observed interactions due to the species-specific nature of the facilitative interactions (Callaway 2007). At the same time, this method allows us to evaluate the contribution of phylogeny and other factors to the structure of the real interaction network (i.e. network connectance and nestedness). Thus, we also tested whether the high levels of nestedness and connectance observed in the real facilitation networks may be explained by both differences in species abundance and phylogenetic history (Verdú and Valiente-Banuet 2008).

Methods

Study site

The study was conducted in the tropical desert community of San Juan Raya (Mexico) located in the northwestern part of the Tehuacán Valley occupying ca 100 km² (18°20′N, 97°28′W). The annual mean temperature is 21°C, and annual rainfall averages about 380 mm, with 85% of it during the summer (Valiente-Banuet et al. 2000). The vegetation is dominated by the columnar cactus *Neobuxbaumia mezzalana* and many shrub species like *Lippia graveolens*, *Calliandra eriophylla*, *Mascagnia seleriana*, *Echinopryx eglandulosa*, *Pseudosmodyngium multifoliolum*, *Acacia subangulata*, *A. constricta*, *Hechtia podantha*, *Cnidosculus tehuacanezis*, *Yucca periculosa* and *Mimosa lacerata*.

Observed matrix of pairwise interactions

The observed matrix of pairwise interactions was constructed by recording the facilitative associations between adult nurses and seedlings of facilitated plants. We sampled four 100 × 10 m² transects to count the number of seedlings (non-reproductive, small plants) of each species growing beneath canopies of adult (reproductive) plants and in open space. Our samplings recorded ca 96% of the whole species inhabiting this plant community. A species was considered to be facilitated when the percentage of individuals recruiting under canopies was greater than expected by the overall canopy cover in the community, estimated by means of the line-interception method as explained below. At the same time we recorded the nurse species with which each seedling was associated to generate a facilitation network. Such network was a matrix containing the number of individuals of each beneficiary species occurring beneath each nurse species. Given that observational and experimental evidence suggest that facilitation in this community occurs between distantly related species (Valiente-Banuet and Verdú 2007, Castillo et al. 2010), we assume that phylogenetic overdispersion underlies the structure of the facilitation network. This contrasts with most phylogenetic analyses that normally test if closely related species tend to behave similarly in the network (Ives and Godfray 2006, Verdú et al. 2010).

Likelihood analysis of pairwise interaction probabilities

We evaluated the ability of abundance and phylogeny to explain the observed matrix of pairwise interactions described above by means of the likelihood approach developed by Vázquez et al. (2009). For comparison with these matrices, we also defined a null probability matrix in which all pairwise interactions were equiprobable.

We calculated the interaction probability matrices expected under the assumptions that interactions were 1) homogeneous across pairs of species (null matrix); 2) determined solely by relative species abundances or phylogeny, and 3) determined by the interaction of abundance and phylogeny matrices.

The null probability matrix was defined as a matrix in which all pairwise interactions had the same probability $1/IJ$ of occurrence, where $I$ and $J$ are the numbers of facilitated and nurse species in the network.

The probability matrix derived of plant abundances was constructed by multiplying the vectors of nurse and facilitated plant abundances. To obtain an estimate of species abundance independent of the data included in the observed matrix of pairwise interactions, we performed different samplings in the same area. As the relevant measure of nurse abundance for facilitated seedlings is nurse cover, we established eight 50 m linear transects in which the coverage of the adult plants was recorded following the line-interception
method (Canfield 1941, Floyd and Anderson 1987). The abundance of facilitated seedlings was estimated by counting the number of seedlings of each species in an extensive sampling of an area of 5000 m$^2$. The underlying assumption of this analysis is that the probability of interaction increases with plant abundances. Because we do not know a priori how estimates of abundances may quantitatively affect the interaction probabilities, we tested different scenarios in which such increasing relationship followed a linear (using the untransformed abundance matrix), power (raising the abundance matrix to a given power) or exponential (exponentiating the abundance matrix curve (see detailed methods in Supplementary material Appendix 1).

The probability matrix derived of plant phylogeny distances was based on the pairwise phylogenetic distances between nurses and facilitated plant species (Fig. 1). The probability of interaction between a nurse and a facilitated species was estimated as the phylogenetic distance between them divided by the sum of all the pairwise distances between all the nurses and all the facilitated species. The phylogenetic distances matrix was obtained from the community phylogeny generated with the help of the program Phylomatic2 and Phylocom 4.1 (Webb et al. 2008). This program generates a community phylogeny by matching the family names of our study species with those contained in a backbone phylogeny, which is the megatree based on the updated work of the Angiosperm Phylogeny Group (Stevens 2005). The megatree was constructed with all phylogenetic information available in the Phylomatic2 repository (accessed 28 January 2011) plus two published phylogenies of Fabaceae (Gómez-Acevedo et al. 2010) and Cactaceae (Hernández-Hernández et al. 2011) containing Mexican species, allowing thus a better resolution of the tree. As the megatree is calibrated with age estimates from Wikstrom et al. (2001), the program returns a calibrated tree with the study species in which the undated nodes have been evenly distributed between dated nodes (Webb et al. 2008). We replaced phylogenetic distance values of zero by $1 \times 10^{-15}$ to avoid infinite values in the likelihood calculation. Different scenarios in which facilitation interactions increased with phylogenetic distances following linear, power or exponential curves were also tested as explained above for abundance (see detailed methods in Supplementary material Appendix 1).

We also calculated combined probabilities of abundance and phylogeny as the element-wise multiplication of abundance and phylogeny matrices. We used the abundance and phylogeny matrices which best performed in the three scenarios (linear, power and exponential) described above (see detailed methods in Supplementary material Appendix 1).

In essence, selecting the probability distributions obtained from the best scenario is equivalent to selecting, say, a log-transformation over other types of data manipulation (e.g. raising the data to a power or performing an arcsine of the square root transformation) when the original data seems to be log-normally distributed.

All the probability matrices were normalized so that their elements added up to one (Supplementary material Appendix 2). We compared the Akaike information criterion (AIC) values across candidate models to select the best fit model. The model with lower AIC value was selected as the best model. As a rule of thumb, models whose AIC is less than 2 units larger than the best fit model also have substantial support, whereas those with models resulting in AIC values $> 10$ units larger have virtually no support (Burnham and Anderson 2002). All the likelihood analyses were run with R-code provided by D. Vázquez.

**Aggregate network statistics**

The ability of the above four probability matrices (null, abundance, phylogeny and abundance × phylogeny) to explain aggregate network statistics as connectance (the proportion of pairs of nurses and facilitated plant species that directly interact) and nestedness (the degree to which specialists interact with proper subsets of the species generalists interact with) was evaluated following Vázquez et al. (2009) approach. We fitted both transformed (exponential

---

**Figure 1.** Derivation of an interaction probability matrix from the phylogenetic relationships between nurses and facilitated species. (A) Phylogenetic tree of facilitated (framed plants) and nurse (unframed plants) species; numbers above branches represent arbitrary units of time spent between two nodes or between a node and a tip; (B) Matrix of phylogenetic distances between nurses (in columns) and facilitated (in rows) species; (C) Matrix of interaction probabilities obtained by dividing each phylogenetic distance by the sum of all the distances.
and power) and untransformed (linear) probability matrices of abundance and phylogeny and the latter always predicted better the network parameters. The method uses a randomization algorithm to evaluate to what extent null, phylogeny and abundance probabilities matrices predicted the observed connectance and nestedness. The algorithm assigns the total number of interactions originally observed in the facilitation matrix according to the four probability matrices, with the only constraint that each species received at least one interaction (see Vázquez et al. 2009 for further mathematical details).

We calculated several measures of nestedness because of the lack of consensus on the best method to define and quantify nestedness (Ulrich et al. 2009). Nestedness was calculated with the BIMMATNEST (Rodrìguez-Gironés and Santamarìa, 2006), NODF (Almeida-Neto et al. 2008) and discrepancy (Brualdi and Sanderson 1999) algorithms. As network metrics that use interaction frequencies instead of presence/absence of interactions have been shown to be more robust against variation in sampling effort (Blüthgen 2010), we also calculated the equivalent of connectance and nestedness for quantitative matrices (Bersier at al. 2002, Tylianakis et al. 2007, Galeano et al. 2009). Quantitative nestedness calculates the nestedness of a network taking into account the weight of the interactions, according to the method proposed by Galeano et al. (2009). Quantitative connectance was calculated as the ratio between the mean number of interactions per species (linkage density) and the number of species in the network (Tylianakis et al. 2007). All the analyses were run with the R-code provided by D. Vázquez modified to include other nestedness and connectance measures as implemented in the bipartite package (Dormann et al. 2009) of R statistical software (R Development Core Team 2009).

**Results**

**Likelihood analysis of pairwise interaction probabilities**

The null matrix was the worst predictor of the observed interaction matrix (Table 1). Phylogeny alone improved the fit of the null matrix but also performed poorly. Both null and phylogeny based matrices alone had AIC’s one order of magnitude larger than that of the observed probability matrix fitted to itself (Table 1). Abundance based models fit better than null and phylogeny models but interestingly the best predictor was by far the matrix combining the abundance and phylogeny probability matrices. In fact, this combined matrix was $(1/1 \times 10^{-53}) = 1 \times 10^{53}$ times more likely to be the best explanation for the observed interaction matrix compared to abundance only (see weights in Table 1). The combined matrix represents the interaction probabilities expected if species interact proportionally to both their phylogenetic distances and abundances. Although this combined matrix was the best predictor, it should be noted that much variation in the observed interactions still remains unexplained as the differences in AIC’s suggest (6020 vs 1088 for the abundance × phylogeny vs observed interaction matrix, Table 1).

**Aggregate network statistics**

The null and phylogeny matrices alone did not predict correctly any of the network statistics observed in the San Juan Raya plant community (Table 2). These matrices strongly overestimated connectance and underestimated nestedness measures.

The abundance matrix correctly predicted all the nestedness values but failed to predict both the qualitative and the quantitative connectance. The combined phylogeny × abundance matrix failed to correctly predict the qualitative measure of connectance but correctly predicted all the nestedness values and the quantitative connectance measure.

**Discussion**

The structure of the real ecological networks is determined by multiple factors including neutral processes as well as factors related to the traits of each species (Vázquez 2005, Rezende et al. 2007). By applying a recent method to evaluate the relative contribution of these factors alone or combined, it becomes immediately evident that single factors cannot explain the variability seen in the ecological interactions among plant species in the Tehuacán system in which we worked. Indeed, our data show that the combined effects of phylogeny and species abundance were, by far, the best predictor of network properties as nestedness and connectance observed in a real plant facilitation network. However, we also find that such information is not enough to correctly predict the occurrence and frequency of pairwise interactions, as reported by Vázquez et al. (2009) in a pollination network. These results pose two questions with interesting consequences. Why is the combined effect of phylogeny and abundance a better predictor of pairwise interactions than phylogeny or abundance alone? Why are we still so far from a correct prediction of the true pairwise interactions?

Obvious phylogeny alone is a bad predictor of the observed ecological interactions because it does not take into account the encounter probability between species, which is mediated by local abundance in the community. Once the encounter probability is accounted for by combining phylogeny and abundance matrices, phylogeny becomes relevant to
explain the observed interactions. On the other side, the fact that the interactions between nurses and facilitated species are worse predicted by abundance alone than by abundance plus phylogeny indicates that this type of interactions is species-specific and therefore that nurses are not replaceable (see Callaway 1998 for a discussion on species-specificity in facilitative interactions). This is an important issue from both theoretical and practical points of view. Theoretically, the identification of highly species-specific facilitative interactions opens the possibility to consider facilitation in local communities as a mechanism of ongoing coevolutionary processes among plants (Thompson 2005, Olesen et al. 2008). This information is also relevant from a conservation point of view if we want to use facilitation as a restoration tool to choose the correct pair of nurse and target plant species (Gómez-Aparicio et al. 2004, Padilla and Pugnaire 2006, Siles et al. 2008).

Table 2. Observed and predicted values of nestedness and connectance in the San Juan Raya plant facilitation network. Predicted values are based on null, abundance and phylogeny matrices. Untransformed matrices of phylogeny and abundance were used for this analysis because of its greater predictive ability relative to transformed matrices. Mean and 95% CI are shown for each predicted value. Estimates correctly predicting the observed value are highlighted in bold.

<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th>Null</th>
<th>Phylogeny</th>
<th>Abundance</th>
<th>Ab × Phylo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connectance</td>
<td>0.21</td>
<td>0.70 [0.68, 0.71]</td>
<td>0.66 [0.64, 0.67]</td>
<td>0.27 [0.26, 0.28]</td>
<td>0.27 [0.25, 0.28]</td>
</tr>
<tr>
<td>Quantitative Connectance</td>
<td>0.122</td>
<td>0.313 [0.305, 0.320]</td>
<td>0.297 [0.290, 0.304]</td>
<td>0.132 [0.127, 0.137]</td>
<td>0.127 [0.122, 0.131]</td>
</tr>
<tr>
<td>Quantitative Nestedness</td>
<td>0.71</td>
<td>0 [0, 0.04]</td>
<td>0.18 [0.14, 0.23]</td>
<td>0.72 [0.67, 0.77]</td>
<td>0.70 [0.65, 0.75]</td>
</tr>
</tbody>
</table>

Why are we still so far from a correct prediction of the true pairwise interactions? Although we have improved substantially the prediction of pairwise interactions by combining abundance and phylogeny, it seems obvious that many more processes are acting simultaneously in shaping such interactions. For example, we can detect some pairwise interactions occurring between closely related species that clearly depart from the expected phylogenetic pattern. These phylogenetically unexpected interactions usually occur between species within Fabaceae (Acacia constricta, Calliandra eryrophylla and Mimosa lacerata facilitating Aescynome compacta). Probably, the ability of Fabaceae species to harbor colonies of nitrogen-fixing bacteria in their roots may compensate their competitive effects (Callaway 2007). Similarly, the negative effects of trees on seedling survival may be offset by the positive effects of mycorrhizal networks established between trees and neighboring seedlings (Booth and Hoeksema 2010). Likewise, indirect effects through third species may be favoring the appearance of the phylogenetically unexpected associations. When many plants grow together indirect effects are usually positive and alleviate direct competitive effects (Callaway 2007, Valiente-Banuet and Verdú 2008). Experimental evidence suggests that the relationships between nurses and beneficiaries in this community are not fully explained by pairwise phylogenetic connections (Castillo et al. 2010). Almost all nurse plants in our system facilitate several species at once, and this creates the potential for very complex indirect and species-specific interactions among different beneficiary species beneath nurses (Cuesta et al. 2010). These indirect interactions are highly likely to modify the pairwise relationships between nurses and beneficiary species that are the basis of our networks. Similarly, Hill and Kotanen (2009) found that the level of herbivory was better predicted by a combined measure of phylogenetic distances of the target plant to all its confamilial species rather than to a single, closest relative, neighbor. Interestingly, these results were only found under controlled – common garden – conditions but not in wild populations, suggesting that phylogenetic influences operate in addition to other sources of ecological variation.

The combination of abundance and phylogeny matrices did not explain all the variation in the observed facilitation interactions but correctly predicted relevant network parameters like nestedness and connectance. Three different causes have been invoked to cause nestedness in interaction matrices: passive sampling, asymmetric interaction strength and phenotypic complementarity (Ulrich et al. 2009). Nestedness through passive sampling is produced when the most abundant species simply have higher chances to interact with their more abundant counterparts. Nestedness through asymmetric interaction strength is produced when some interactions are ‘forbidden’ because of biological constraints, such as phenological asynchrony or morphological mismatching between the interacting species. Paradoxically, phenotypic complementarity, the opposite extreme to morphological mismatching, may also cause nestedness. Rezende et al. (2007) demonstrated that complementarity between phytophagous traits of plants and animals can explain the nested pattern of interaction networks, particularly when several stages are involved. As phenotypes are complex arrays of multiple traits, it is very difficult to properly characterize the complementarity between multidimensional phenotypes of species. Instead, the phylogenetic relationship between species may inform us about phenotype complementarity given that closely related species tend to be similar in several phytophagous traits because they inherit multiple traits from a single ancestor. This is exactly the case for the morphological and physiological traits defining the regeneration niche of plants, which are evolutionarily conserved and therefore closely related species tend to have the same regeneration niche (Valiente-Banuet et al. 2006, Valiente-Banuet and Verdú 2007, Verdú and Pausas 2007, Pausas and Verdú 2008). Accordingly, facilitative interactions are produced by connecting distantly-related species with complementary regeneration niches (nurses and facilitated plants) and this pattern of interactions also produces significant nestedness.

In short, we have shown that the phylogenetic history, acting together with other ecological factors, has a pervasive influence in the structure of ecological networks.
Acknowledgements – We are very grateful to Diego Vázquez for providing the R-code and helping with the analyses. A. Vital, C. Silva and J. P. Castillo helped with field samplings. R. M. Callaway, M. Goberna, E. Rezende and D. Vázquez provided valuable comments on the manuscript. This work was funded by AECID (projects A017475/08, A023461/09), DGAPA-UNAM (project IN-224808-3), and CYTED (Acción 409AC0369).

References
Cuesta, B. et al. 2010. Facilitation of Quercus ilex in Mediterranean shrubland is explained by both direct and indirect interactions mediated by herbs. – J. Ecol 98: 687–696.

Supplementary material (available online as Appendix O19477 at <www.oikosoffice.lu.se/appendix>). Appendix 1–2.