

# Evaluating multiple determinants of the structure of plant–animal mutualistic networks\*

Diego P. Vázquez<sup>1,2</sup>, Natacha P. Chacoff<sup>1</sup> and Luciano Cagnolo<sup>3</sup>

1. Instituto Argentino de Investigaciones de las Zonas Áridas, CONICET, CC 507, 5500 Mendoza, Argentina (dvazquez@mendoza-conicet.gov.ar, nchacoff@mendoza-conicet.gov.ar)
2. Instituto de Ciencias Básicas, Universidad Nacional de Cuyo, Centro Universitario, M5502JMA Mendoza, Argentina
3. IMBIV and Centro de Investigaciones Entomológicas de Córdoba, CONICET–Universidad Nacional de Córdoba, FCEF N, Av. Vélez Sarsfield 1611, X5016GCA Córdoba, Argentina (lcagnolo@efn.uncor.edu)

**Running title:** The structure of mutualistic networks

**Manuscript type:** Report

## Abstract

The structure of mutualistic networks is likely to result from the simultaneous influence of neutrality and the constraints imposed by complementarity in species phenotypes, phenologies, spatial distributions, phylogenetic relationships and sampling artifacts. We develop a conceptual and methodological framework to evaluate the relative contributions of these potential determinants. Applying this approach to the analysis of a plant–pollinator network, we show that information on relative abundance and phenology suffices to predict several aggregate network properties (connectance, nestedness, interaction evenness and interaction asymmetry). However, such information falls short of predicting the detailed network structure (the frequency of pairwise interactions), leaving a large amount of variation unexplained. Taken together, our results suggest that both relative species abundance and complementarity in spatio–temporal distribution contribute substantially to generate observed network patterns, but that this information is by no means sufficient to predict the occurrence and frequency of pairwise interactions. Future studies could use our methodological framework to evaluate the generality of our findings in a representative sample of study systems with contrasting ecological conditions.

*Keywords: forbidden links; neutrality; phenotypic complementarity; phylogenetic signal; pollination; Villavicencio network*

---

\*Article in press in *Ecology*

## 34 Introduction

35 There is growing interest in the study of networks of interacting plant and animal  
 36 mutualists. This interest stems from the realization that considering the community  
 37 context is important to understand the ecological and evolutionary implications of  
 38 mutualistic interactions (Strauss and Irwin, 2004). Studies of mutualistic networks have  
 39 uncovered some apparently general structural properties, such as the skewed distribution  
 40 of links per species (many species with few links and few species with many links;  
 41 Jordano et al., 2003; Vázquez and Aizen, 2003), the nested organization of the interaction  
 42 matrix (Bascompte et al., 2003) and the frequent occurrence of asymmetric interactions  
 43 (Vázquez and Aizen, 2004; Bascompte et al., 2006).

44 Recent discussion about the potential ecological and evolutionary determinants of  
 45 these structural patterns has centered around the relative importance of neutrality vs.  
 46 the so-called “forbidden links”. The neutrality hypothesis posits that network patterns  
 47 result from the fact that individuals interact randomly, so that abundant species interact  
 48 more frequently and with more species than rare species (Dupont et al., 2003; Ollerton  
 49 et al., 2003; Vázquez et al., 2007). The forbidden links hypothesis posits that network  
 50 patterns result from constraints to interactions imposed by the complementarity in  
 51 species phenotypes, phenologies, spatial distributions and phylogenetic relationships  
 52 (Jordano et al., 2003; Rezende et al., 2007; Santamaría and Rodríguez-Gironés, 2007;  
 53 Stang et al., 2007). For example, two species cannot interact if their phenologies do not  
 54 overlap, regardless of what their abundance alone predicts.

55 Available evidence suggests that both neutrality and forbidden links contribute to some  
 56 extent to determine network structure (Bascompte and Jordano, 2007; Vázquez et al.,  
 57 2009). For example, Vázquez et al. (2007) have shown that relative species abundance  
 58 partly (but not entirely) explains the observed asymmetry in the strength of pairwise  
 59 interactions, whereas Stang et al. (2007) showed that information on both abundance and  
 60 morphological traits of plants and pollinators are needed to predict asymmetry and  
 61 nestedness in binary networks. Several recent studies have also shown that phenologies  
 62 and inter-annual dynamics of plant and animal species influence network structure  
 63 (Basilio et al., 2006; Alarcón et al., 2008; Olesen et al., 2008; Petanidou et al., 2008). In a  
 64 similar vein, a recent simulation study has shown that the degree of mixing resulting from  
 65 the spatial aggregation of plant individuals and the scale of animal movement decisions  
 66 has strong influences on network structure (Morales and Vázquez, 2008). The key  
 67 unanswered question is how important each of these processes is in determining network  
 68 structure. Here we develop a conceptual and methodological framework to answer this  
 69 question and apply it to investigate the determinants of a plant–pollinator network.

70 Consider a mutualistic network depicted as an interaction matrix  $\mathbf{Y}$  with  $I$  rows and  $J$   
 71 columns corresponding to the plant and animal species in the network, respectively, and a  
 72 positive integer in cell  $y_{ij}$  representing the number of interactions recorded between plant  
 73  $i$  and animal  $j$ . This matrix is a function of multiple interaction probability matrices of  
 74 the same size as  $\mathbf{Y}$ , determined by relative species abundance ( $\mathbf{N}$ ), temporal ( $\mathbf{T}$ ) and  
 75 spatial overlap ( $\mathbf{S}$ ) and phenotypic traits of interacting species ( $\mathbf{K}$ ). The effects of these  
 76 factors on  $\mathbf{Y}$  can be constrained by the phylogenetic relationships among plants ( $\mathbf{P}_p$ ) and  
 77 animals ( $\mathbf{P}_a$ ) (Rezende et al., 2007). In addition, detection probabilities of interactions  
 78 resulting from sampling effects ( $\mathbf{E}$ ) can also influence the observed network (Blüthgen  
 79 et al., 2008). Thus,

$$\mathbf{Y} = f(\mathbf{N}, \mathbf{T}, \mathbf{S}, \mathbf{K}, \mathbf{P}_p, \mathbf{P}_a, \mathbf{E}). \quad (1)$$

80 Below we use this conceptual framework to evaluate the contribution of abundance,  
 81 spatial and temporal overlap and phylogenetic relatedness among species on the structure  
 82 of a plant–pollinator network. Specifically, we address the following questions: (i) To  
 83 what extent do relative abundance and spatio–temporal overlap predict aggregate  
 84 network statistics (connectance, nestedness, interaction strength evenness and the  
 85 distribution of interaction strength asymmetries)? (ii) To what extent do these factors  
 86 predict pairwise interaction frequencies in the interaction matrix? (iii) Is there any  
 87 detectable phylogenetic signal in the interaction matrix, suggesting that the influence of  
 88 abundance and spatio–temporal overlap on network structure could have resulted from  
 89 phylogenetic constraints imposed by the phylogenetic relationships among plants and  
 90 among animals?

## 91 Materials and Methods

92 **Study system** Data come from a plant–pollinator network from the Monte desert at  
 93 Villavicencio Nature Reserve (32° 32' S, 68° 57' W, 1270 masl), Mendoza, Argentina. We  
 94 worked in four 1 ha plots, separated by 1–2 km. Predominant vegetation is a tall  
 95 shrubland dominated by *Larrea divaricata*, *Zuccagnia punctata*, *Prosopis flexuosa*,  
 96 *Condalia microphylla*, *Acantholippia seriphoides* and *Opuntia sulphurea* (Roig, 1972). We  
 97 give only a summarized description of field methods here; further details can be found in  
 98 the original publication describing the network (Chacoff et al., 2009).

99 **Plant–pollinator interactions** Flower visiting insects were observed on plant species  
 100 in weekly surveys in two consecutive flowering seasons (2006 and 2007) between  
 101 September and January (2006) or December (2007). We attempted to sample  
 102 plant–pollinator interactions in the whole community as comprehensively as possible,  
 103 recording interactions between 41 plant species and 97 insect species. With these data we  
 104 constructed a quantitative plant–pollinator interaction matrix  $\mathbf{Y} = [y_{ij}]$ , with rows  
 105 corresponding to plant species and columns to pollinator species; cell entries  $y_{ij}$  are  
 106 integers representing the number of flowers of plant species  $i$  visited by pollinator species  
 107  $j$  (Fig. 1a). This is the network we want to predict.

108 **Species abundance** Plant abundance was assessed in weekly surveys along five fixed  
 109 50 m × 2 m transects in the four sites, where we recorded the number of individuals of  
 110 each entomophilous species and, for a subset of individuals, the number of flowers per  
 111 individual. We also collected three fresh flowers from ten individuals of each species to  
 112 estimate the number of pollen grains produced per flower. We attempted to obtain nectar  
 113 from flowers, but we failed for most plant species; flowers in this system usually have very  
 114 small standing volumes of nectar. With these data we estimated the density of  
 115 individuals, density of flowers (density of individuals × flowers per individual) and  
 116 density of pollen (density of individuals × flowers per individual × pollen grains per  
 117 flower) for each species at each site. Arguably, density of flowers is the most appropriate  
 118 measure of abundance from the flower visitors' perspective: it is a better estimate of  
 119 resource abundance than density of individual plants (because there is high variation  
 120 among plant species in the mean number of flowers per individual) and focuses on the  
 121 flower as the resource unit for both pollen and nectar. However, because different studies  
 122 use different measures of abundance, we considered the three measures to evaluate how  
 123 the choice of a particular measure affects our results.

124 Because flowers of most species usually last less than a week, and because we were  
 125 interested in an overall measure of abundance, we summed density across weeks and sites  
 126 to obtain an overall estimate of abundance of each species. Thus, a plant species could be  
 127 abundant because it produced many flowers in a short period or because it produced few  
 128 flowers over a long period. These two forms of abundance were then distinguished by  
 129 incorporating temporal and spatial structure in our network model, as described below  
 130 (see Spatial and temporal overlap).

131 Insect abundance was defined as the total number of visits made by a particular insect  
 132 species to any plant species (that is, the column sum of the interaction matrix), as done  
 133 in previous studies (see, e.g., Vázquez et al., 2007).

134 **Spatial and temporal overlap** To quantify spatial and temporal overlap of plant and  
 135 pollinator species, we first compiled matrices of temporal and spatial occurrences, with  
 136 species in rows and date or site in columns, and cells with ones for presences and zeros for  
 137 absences. We thus had one temporal and one spatial occurrence matrix for plants,  $\mathbf{O}_p^t$   
 138 and  $\mathbf{O}_p^s$ , and one of each for animals,  $\mathbf{O}_a^t$  and  $\mathbf{O}_a^s$ . We then used matrix multiplication to  
 139 calculate spatial and temporal overlap matrices between plants and animals,  $\mathbf{S} = \mathbf{O}_p^s \mathbf{O}_a^{s'}$   
 140 and  $\mathbf{T} = \mathbf{O}_p^t \mathbf{O}_a^{t'}$ , where the prime symbol indicates the transpose of a matrix or vector.

141 **Calculation of interaction probabilities** We calculated interaction probability  
 142 matrices expected under the assumptions that interactions were determined solely by  
 143 relative species abundances, temporal overlap and spatial overlap. For relative abundance,  
 144 interaction probability between a plant species  $i$  and a pollinator species  $j$  is simply the  
 145 product of their relative abundances. In matrix notation, the interaction probability  
 146 matrix expected from relative abundances is  $\mathbf{N} = \mathbf{n}_p \mathbf{n}_a'$ . For temporal and spatial  
 147 overlap, we used overlap matrices  $\mathbf{T}$  and  $\mathbf{S}$  normalized so that their elements added up to  
 148 one, so as to transform them into probabilities. (For simplicity, we call these normalized  
 149 matrices  $\mathbf{T}$  and  $\mathbf{S}$  hereafter.) Thus, the greater the temporal or spatial overlap of two  
 150 species, the greater their probability of interaction; species with no temporal or spatial  
 151 overlap had zero probability of interaction. We also calculated combined probabilities as  
 152 the element-wise multiplication of matrices  $\mathbf{N}$ ,  $\mathbf{T}$  and  $\mathbf{S}$ , again normalizing the resulting  
 153 matrices so that their elements added up to one. These combined matrices represent the  
 154 expected probability under the joint influence of more than one of these factors. For  
 155 example,  $\overline{\mathbf{NS}}$  denotes the combined abundance–spatial overlap probability matrix, and  
 156 represents the interaction probabilities expected if species interact proportionally to their  
 157 abundances given that they co-occur at a particular site. Thus, we had seven probability  
 158 matrices with all possible combinations of relative abundance and temporal and spatial  
 159 overlap:  $\mathbf{N}$ ,  $\mathbf{T}$ ,  $\mathbf{S}$ ,  $\overline{\mathbf{NT}}$ ,  $\overline{\mathbf{NS}}$ ,  $\overline{\mathbf{TS}}$  and  $\overline{\mathbf{NTS}}$ . In addition, we defined an eighth probability  
 160 matrix in which all pairwise interactions had the same probability  $1/IJ$  of occurrence,  
 161 where  $I$  and  $J$  are the numbers of plant and animal species in the network; this  
 162 probability matrix, termed “Null” below (see Figs. 2 and 3), is taken as a benchmark null  
 163 model for comparison with the other seven probability matrices.

164 **Analysis of aggregate network statistics** We considered four aggregate network  
 165 statistics frequently used in the analysis of plant–animal mutualistic networks:  
 166 connectance, nestedness, interaction evenness and interaction asymmetry. Connectance is  
 167 the proportion of realized interspecific links, defined as  $C = L/(IJ)$ , where  $L$  is the  
 168 number of non-zero entries in the binary interaction network and  $I$  and  $J$  are, as above,

169 the numbers of plant and animal species in the network. Nestedness is the tendency of  
 170 specialized species to interact with a subset of the interaction partners of more  
 171 generalized species or, more precisely, the degree of symmetry in the distribution of  
 172 unexpected absences and presences on each side of the boundary line defining perfect  
 173 nestedness (Almeida-Neto et al., 2007). Thus, a nestedness value of 1 represents complete  
 174 lack of symmetry (perfect nestedness), while a value of 0 represents the highest symmetry  
 175 in the distribution of unexpected presences and absences, with absolute randomness  
 176 falling somewhere in between 0 and 1. Nestedness was calculated with the  
 177 BINMATNEST algorithm proposed by Rodríguez-Gironés and Santamaría (2006),  
 178 implemented in the bipartite package (Dormann et al., 2008) of R statistical software (R  
 179 Development Core Team, 2007); the BINMATNEST algorithm overcomes several  
 180 limitations of the widely used nestedness temperature calculator developed by Atmar and  
 181 Patterson (1993). Following Tylianakis et al. (2007), interaction evenness was defined as  
 182 Shannon’s index,  $H = p_{ij} \log_2 p_{ij} / \log_2 F$ , where  $F$  is the total number of plant–pollinator  
 183 interactions in the matrix (see eq. 2 below) and  $p_{ij}$  is the proportion of those interactions  
 184 involving plant  $i$  and pollinator  $j$ . An uneven network is one with high skewness in the  
 185 distribution of interaction frequencies. Interaction asymmetry for a given species was  
 186 defined as  $A_i = \frac{\sum_j d_{ij}}{k_i}$ , where  $k_i$  is the degree of species  $i$  (i.e., the number of species with  
 187 which  $i$  interacts) and  $d_{ij}$  is a measure of the symmetry of the strength of the pairwise  
 188 interaction between  $i$  and  $j$  (Vázquez et al., 2007); as in previous publications, we used  
 189 interaction frequency as a surrogate of interaction strength (see Vázquez et al., 2005;  
 190 Bascompte et al., 2006; Vázquez et al., 2007).

191 We used a randomization algorithm implemented in R (see Supplement) to evaluate to  
 192 what extent interaction probabilities derived from relative abundance and spatial and  
 193 temporal overlap of species occurrences predicted the observed aggregate network  
 194 statistics. The algorithm assigned the total number of interactions originally observed in  
 195 the interaction matrix according to the seven probability matrices defined by all possible  
 196 combinations of abundance and temporal and spatial overlap (see Interaction probabilities  
 197 above), with the only constraint that each species received at least one interaction.

198 **Likelihood analysis of pairwise interaction probabilities** We used a likelihood  
 199 approach to evaluate the ability of abundance, temporal overlap and spatial overlap to  
 200 predict the detailed structure of the interaction matrix. Consider the observed interaction  
 201 matrix  $\mathbf{Y}$  and a probability matrix  $\mathbf{X}$ , whose entries  $x_{ij}$  are the probabilities of  
 202 occurrence for each pairwise interaction; we want to evaluate whether those probabilities  
 203 match the observed frequencies of interaction. This evaluation can be done by calculating  
 204 the likelihood of probability matrix  $\mathbf{X}$  given the data ( $\mathbf{Y}$ ). We assumed that the pairwise  
 205 probability of interaction between a plant  $i$  and a pollinator  $j$  followed a multinomial  
 206 distribution. Thus, the likelihood of probability matrix  $l$  given the data is

$$\mathcal{L}_l = \frac{F!}{\prod_{i=1}^I \prod_{j=1}^J y_{i,j}!} \prod_{i=1}^I \prod_{j=1}^J x_{ij}^{y_{ij}} \quad (2)$$

207 where  $F$  is the total number of observed interactions (i.e., the sum of the elements of  
 208 matrix  $\mathbf{Y}$ ), and  $I$  and  $J$  are the total number of animals and plants in  $\mathbf{Y}$ , respectively.  
 209 We calculated this likelihood using function `dmultinom` in the `stats` package of R. We  
 210 then calculated Akaike’s Information Criterion as  $AIC_l = L_l - 2k_l$ , where  $k_l$  is the  
 211 number of parameters involved in generating probability matrix  $l$ . The number of

212 parameters was defined as the number of factors contributing to generate a particular  
 213 probability matrix; thus, matrix  $\mathbf{N}$  has one parameter,  $\overline{\mathbf{TS}}$  has two and  $\overline{\mathbf{NTS}}$  has three.

214 **Phylogenetic analysis** Ideally, we would like to incorporate phylogenetic effects into  
 215 our conceptual framework, constructing a phylogenetic probability matrix as we did with  
 216 abundance, temporal overlap and spatial overlap. Regrettably, incorporation of  
 217 phylogenetic effects was not possible because we found no way of predicting pairwise  
 218 interactions from the plant and pollinator phylogenies alone, without reference to the  
 219 observed interaction matrix. Alternatively, we evaluated the strength of the phylogenetic  
 220 signal of the two phylogenies on the interaction matrix with the method developed by  
 221 Ives and Godfray (2006): a linear model approach to fit the phylogenetic  
 222 variance–covariance matrix to the interaction matrix. Using this method we calculated  
 223 the independent signals of the plant ( $d_p$ ) and animal ( $d_a$ ) phylogenies and the strength of  
 224 the signal of both phylogenies combined ( $\text{MSE}_d$ ) (Appendix A).

## 225 Results

226 **Aggregate network statistics** No probability matrix predicted connectance and  
 227 interaction evenness values whose confidence intervals included observed values of these  
 228 statistics (Fig. 2a, c). However, predictions from probability matrices  $\overline{\mathbf{NTS}}$  and  $\overline{\mathbf{NT}}$ ,  $\overline{\mathbf{NS}}$   
 229 and  $\mathbf{N}$  were extremely close to observed connectance and interaction evenness. Observed  
 230 nestedness was included within the randomization confidence intervals of the same four  
 231 probability matrices (Fig. 2b). A similar result was observed for interaction asymmetry  
 232 for pollinators (Fig. 2d), but not for plants (Fig. 2e). In the latter case, observed average  
 233 asymmetry was close to zero (predominantly symmetric interactions) and similar to that  
 234 predicted by the null model that assumes that all species have the same probability of  
 235 interaction. Highly negative asymmetry in pollinators and high symmetry for plants  
 236 matches previous results for other interaction networks (Vázquez et al., 2007). Thus,  
 237 with the exception of asymmetry for plants, information on abundance and temporal  
 238 overlap suffices to simulate interaction networks whose aggregate structure resembles very  
 239 closely the structure of the observed matrix  $\mathbf{Y}$ . This result was the same regardless of the  
 240 measure of abundance used (i.e., density of individuals, individuals  $\times$  flowers or  
 241 individuals  $\times$  flowers  $\times$  pollen; result not shown).

242 **Frequency of pairwise interactions** Taking density of flowers as our measure of  
 243 plant abundance (which, as explained above, is arguably the most appropriate measure),  
 244 the combined probability matrix  $\overline{\mathbf{NT}}$  had the lowest dAIC value, with a difference with  
 245 the next best-fitting probability matrix  $\overline{\mathbf{TS}}$  of 156, and several orders of magnitude with  
 246 the null model (Fig. 3b). However, this best-fitting probability model was also orders of  
 247 magnitude worse than the perfect fit obtained by fitting the interaction matrix to itself.  
 248 These results suggest that abundance and temporal overlap are useful to predict part of  
 249 the detailed structure of the interaction matrix, but that we are far from an accurate  
 250 prediction, with much unexplained variation.

251 The latter conclusion can be visualized by comparing the observed interaction matrix  
 252 (Fig. 1a) with an example of a matrix resulting from the randomization procedure with  
 253 the  $\overline{\mathbf{NT}}$  probability matrix (Fig. 1b). For example, although the two randomized  
 254 matrices look roughly similar to the observed matrix in terms of connectance and  
 255 nestedness, it is obvious that the most frequently observed interactions are not those

256 predicted to occur most frequently by  $\overline{\mathbf{NT}}$ . Thus, although knowledge of abundance and  
 257 phenology allows us to predict aggregate network properties with high accuracy, we do  
 258 rather poorly at predicting the detailed occurrence of pairwise interactions.

259 Unlike results for aggregate network statistics, using a different measure of abundance  
 260 affects our ability to predict pairwise interactions (Figs. 3a, c):  $\overline{\mathbf{NT}}$  is now in sixth place,  
 261 doing particularly badly when abundance is measured as density of pollen, almost as  
 262 badly as the null model. Notice that because matrices not including abundance are the  
 263 same between Figs. 3a, b and c, results can be directly compared, indicating that the  
 264  $\overline{\mathbf{NT}}$  probability matrix when abundance is measured as density of flowers provides the  
 265 overall best fit.

266 **Phylogenetic signal** The independent phylogenetic signal of the insect phylogeny was  
 267 weak ( $d_a = 0.067$ ) and its confidence interval overlapped zero (95% confidence limits: [0,  
 268 0.286]). The independent plant phylogenetic signal was stronger ( $d_p = 0.327$ ) and did not  
 269 overlap zero (95% confidence limits: [0.189, 0.532]). The overall strength of the  
 270 phylogenetic signal for the linear model fitted to the actual data ( $\text{MSE}_d = 250.05$ ) was  
 271 much closer to that found under the assumption of no phylogenetic covariances (the “star”  
 272 phylogeny:  $\text{MSE}_{\text{star}} = 269.29$ ) than for the assumption of maximum phylogenetic signal  
 273 (Brownian motion evolution:  $\text{MSE}_b = 420.61$ ). Taken together, these results suggest that  
 274 only phylogenetic relationships among plants, not insects, impose some structure on the  
 275 interaction matrix  $\mathbf{Y}$ , but that the overall phylogenetic signal is extremely weak.

## 276 Discussion

277 We have developed a conceptual and methodological framework to evaluate the  
 278 simultaneous contributions of neutrality, forbidden links and sampling effects on the  
 279 structure of mutualistic networks. We have used this framework to evaluate the influence  
 280 of relative abundance (neutrality) and spatio-temporal overlap on a pollination network.  
 281 Although we have not included information on phenotypic trait matching and sampling  
 282 effects, if available this type of information can be easily incorporated to our conceptual  
 283 and methodological framework. For example, Stang et al. (2006; 2007; 2009) have  
 284 derived interaction probabilities based on rules of phenotypic matching between plants  
 285 and nectar-feeding flower visitors. From such information it would be straightforward to  
 286 derive a probability matrix to evaluate the relative contribution of phenotypic matching  
 287 to the structure of the interaction matrix. Similarly, if detection probabilities of  
 288 particular interactions could be estimated, then a detection probability matrix  $\mathbf{E}$  could  
 289 be incorporated into our framework (our implicit assumption above has been that all  
 290 interactions have a detection probability of one).

291 One issue we have not been able to solve is how to incorporate phylogenetic  
 292 information into our framework. This limitation stems from the difficulty of predicting  
 293 interaction probabilities based on the independent plant and animal phylogenies, with no  
 294 reference to the observed interaction matrix. Current methods for the detection of a  
 295 phylogenetic signal in interaction networks (Ives and Godfray, 2006) use a linear model  
 296 approach, fitting the phylogenetic variance-covariance matrix to the interaction matrix,  
 297 and there is currently no way of deriving an expected probability matrix based on the  
 298 independent phylogenies alone. We hope our efforts can stimulate others to work out a  
 299 solution to this crucial problem.

300 Using our conceptual and methodological framework, we have shown that interaction  
 301 probabilities derived from abundance and temporal overlap predict very closely the  
 302 aggregate properties of a plant–pollinator network. In contrast, our likelihood analysis  
 303 shows that information on abundance and spatio–temporal distribution falls short of  
 304 predicting the detailed network structure, leaving a large amount of variation  
 305 unexplained. Thus, although information on abundance and spatio–temporal overlap  
 306 allowed us to construct networks with the same aggregate features of real-world networks,  
 307 we failed resoundingly when attempting to delve into the details of pairwise species  
 308 interactions, which is arguably the ecologically and evolutionarily relevant scale of  
 309 analysis. Taken together, our results suggest that both relative species abundance and  
 310 complementarity in spatio–temporal distribution contribute substantially to generate  
 311 observed network patterns, but that this information is by no means sufficient to predict  
 312 the detailed structure of the interaction network.

313 Of course, the above results for the Villavicencio network are by no means a general  
 314 evaluation of the relative contribution of neutrality and forbidden links to the structure of  
 315 mutualistic networks. Only future studies applying our (or a similar) approach to  
 316 multiple datasets will allow such general evaluation. Unfortunately, most datasets  
 317 available to date do not include the sort of detailed information needed for this  
 318 comparison. Clearly, further progress in the understanding of the determinants of  
 319 network patterns requires spatio–temporally explicit datasets with detailed natural  
 320 history information that may allow deriving sensible rules of phenotypic complementarity.  
 321 We believe this goal will be facilitated if research efforts are focused on a sample of  
 322 representative study systems with contrasting ecological conditions throughout the world.

## 323 Acknowledgments

324 NPC and DPV are career researchers and LC a post-doctoral fellow with CONICET.  
 325 Research was funded through grants from CONICET (PIP 6564), FONCYT–ANPCYT  
 326 (PICT 20805) and BBVA Foundation (BIOCON03-162). The following colleagues helped  
 327 with taxonomic identifications: Federico Agrain (coleopterans), Cecilia Domínguez  
 328 (dipterans), Roberto Kiesling, Eduardo Martínez Carretero and Eduardo Méndez  
 329 (plants), and Arturo Roig-Alsina (bees). We thank Federico Agrain, Mathew Helmus,  
 330 Anthony Ives, Juan Manuel Morales, William Morris and Gunilla Sthåls for discussion  
 331 and Nico Blüthgen and Silvia Lomáscolo for comments on the manuscript.

## 332 References

- 333 Alarcón, R., N. M. Waser, and J. Ollerton. 2008. Year-to-year variation in the topology  
 334 of a plant–pollinator interaction network. *Oikos* **117**:1796–1807.
- 335 Almeida-Neto, M., P. R. Guimarães Jr, and T. M. Lewinsohn. 2007. On nestedness  
 336 analyses: rethinking matrix temperature and anti-nestedness. *Oikos* **116**:716–722.
- 337 Atmar, W. and B. D. Patterson. 1993. The measure of order and disorder in the  
 338 distribution of species in fragmented habitat. *Oecologia* **93**:373–382.
- 339 Bascompte, J. and P. Jordano. 2007. Plant–animal mutualistic networks: the architecture  
 340 of biodiversity. *Annual Review of Ecology, Evolution and Systematics* **38**:567–593.

- 341 Bascompte, J., P. Jordano, C. J. Melián, and J. M. Olesen. 2003. The nested assembly of  
342 plant-animal mutualistic networks. *Proceedings of the National Academy of Sciences*  
343 USA **100**:9383–9387.
- 344 Bascompte, J., P. Jordano, and J. M. Olesen. 2006. Asymmetric coevolutionary networks  
345 facilitate biodiversity maintenance. *Science* **312**:431–433.
- 346 Basilio, A. M., D. Medan, J. P. Torretta, and N. J. Bartoloni. 2006. A year-long  
347 plant–pollinator network. *Austral Ecology* **31**:975–983.
- 348 Blüthgen, N., J. Fründ, D. P. Vázquez, and F. Menzel. 2008. What do interaction  
349 network metrics tell us about specialization and biological traits? *Ecology*  
350 **89**:3387–3399.
- 351 Chacoff, N. P., D. P. Vázquez, E. L. Stevani, J. Dorado, and B. Padrón. 2009. The  
352 structure of a desert plant–pollinator network. In review .
- 353 Dormann, C. F., B. Gruber, and J. Fründ. 2008. The bipartite Package, version 0.5. R  
354 Project for Statistical Computing.
- 355 Dupont, Y. L., D. M. Hansen, and J. M. Olesen. 2003. Structure of a plant-flower-visitor  
356 network in the high-altitude sub-alpine desert of Tenerife, Canary Islands. *Ecography*  
357 **26**:301–310.
- 358 Ives, A. R. and H. C. J. Godfray. 2006. Phylogenetic analysis of trophic associations.  
359 *American Naturalist* **168**:E1–E14.
- 360 Jordano, P., J. Bascompte, and J. M. Olesen. 2003. Invariant properties in  
361 coevolutionary networks of plant–animal interactions. *Ecology Letters* **6**:69–81.
- 362 Morales, J. M. and D. P. Vázquez. 2008. The effect of space in plant-animal mutualistic  
363 networks: insights from a simulation study. *Oikos* **117**:1362–1370.
- 364 Olesen, J. M., J. Bascompte, H. Elberling, and P. Jordano. 2008. Temporal dynamics in  
365 a pollination network. *Ecology* **89**:1573–1582.
- 366 Ollerton, J., S. D. Johnson, L. Cranmer, and S. Kellie. 2003. The pollination ecology of  
367 an assemblage of grassland asclepiads in south africa. *Annals of Botany* **92**:807–834.
- 368 Petanidou, T., A. S. Kallimanis, J. Tzanopoulos, S. P. Sgardelis, and J. D. Pantis. 2008.  
369 Long-term observation of a pollination network: fluctuation in species and interactions,  
370 relative invariance of network structure and implications for estimates of specialization.  
371 *Ecology Letters* **11**:564–575.
- 372 R Development Core Team, 2007. R: A Language and Environment for Statistical  
373 Computing. R Foundation for Statistical Computing, Vienna, Austria. URL  
374 <http://www.R-project.org>.
- 375 Rezende, E. L., J. E. Lavabre, P. R. Guimarães, P. Jordano, and J. Bascompte. 2007.  
376 Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature*  
377 **448**:925–928.

- 378 Rodríguez-Gironés, M. A. and L. Santamaría. 2006. A new algorithm to calculate the  
379 nestedness temperature of presence–absence matrices. *Journal of Biogeography*  
380 **33**:924–935.
- 381 Roig, F. A. 1972. Bosquejo fisionómico de la vegetación de la provincia de mendoza.  
382 *Boletín de la Sociedad Argentina de Botánica* **13**, **Suppl.**:49–80.
- 383 Santamaría, L. and M. A. Rodríguez-Gironés. 2007. Linkage rules for plant–pollinator  
384 networks: trait complementarity or exploitation barriers? *PLoS Biology* **5**:e31.
- 385 Stang, M., P. G. L. Klinkhamer, and E. van der Meijden. 2006. Size constraints and  
386 flower abundance determine the number of interactions in a plant–flower visitor web.  
387 *Oikos* **112**:111–121.
- 388 Stang, M., P. G. L. Klinkhamer, and E. van der Meijden. 2007. Asymmetric  
389 specialization and extinction risk in plant–flower visitor webs: a matter of morphology  
390 or abundance? *Oecologia* **151**:442–453.
- 391 Stang, M., P. G. L. Klinkhamer, N. M. Waser, I. Stang, and E. van der Meijden. 2009.  
392 Size-specific interaction patterns and size matching in a plant–pollinator interaction  
393 web. *Annals of Botany* **103**:000–000.
- 394 Strauss, S. Y. and R. E. Irwin. 2004. Ecological and evolutionary consequences of  
395 multispecies plant–animal interactions. *Annual Review of Ecology, Evolution, and*  
396 *Systematics* **35**:435–466.
- 397 Tylianakis, J. M., T. Tschardt, and O. T. Lewis. 2007. Habitat modification alters the  
398 structure of tropical host–parasitoid food webs. *Nature* **445**:202–205.
- 399 Vázquez, D. P. and M. A. Aizen. 2003. Null model analyses of specialization in  
400 plant–pollinator interactions. *Ecology* **84**:2493–2501.
- 401 Vázquez, D. P. and M. A. Aizen. 2004. Asymmetric specialization: a pervasive feature of  
402 plant–pollinator interactions. *Ecology* **85**:1251–1257.
- 403 Vázquez, D. P., N. Blüthgen, L. Cagnolo, and N. P. Chacoff. 2009. Uniting pattern and  
404 process in plant–animal mutualistic networks: a review. *Annals of Botany* **103**:000–000.
- 405 Vázquez, D. P., C. J. Melián, N. M. Williams, N. Blüthgen, B. R. Krasnov, and  
406 R. Poulin. 2007. Species abundance and asymmetric interaction strength in ecological  
407 networks. *Oikos* **116**:1120–1127.
- 408 Vázquez, D. P., W. F. Morris, and P. Jordano. 2005. Interaction frequency as a surrogate  
409 for the total effect of animal mutualists on plants. *Ecology Letters* **8**:1088–1094.

## 410 Figure legends

411 **Figure 1.** Plant–pollinator interaction matrices. (a) Observed plant–pollinator matrix  
 412 in the Monte desert of Villavicencio Nature Reserve (Mendoza, Argentina). (b)  
 413 Interaction matrix resulting from one iteration of the randomization algorithm, using the  
 414  $\overline{\mathbf{NT}}$  probability matrix to assign interactions. In each matrix, rows represent pollinator  
 415 species, columns represent plant species, and circle diameter of a matrix element  $y_{ij}$  is  
 416 proportional to the square root of the number of interactions between pollinator  $i$  and  
 417 plant  $j$ .

418 **Figure 2.** Comparison between aggregate network statistics observed in the  
 419 Villavicencio network and those predicted by the probability matrices. In each panel, the  
 420 vertical line represents the observed value of an aggregate statistic, the circles represent  
 421 the value of the statistic expected from each probability matrix, with errorbars indicating  
 422 95% confidence intervals. Results are shown for the seven probability matrices resulting  
 423 from all possible combinations of abundance (**N**), temporal overlap (**T**) and spatial  
 424 overlap (**S**) and for the null probability matrix with homogeneous interaction  
 425 probabilities across all pairwise interactions (Null).

426 **Figure 3.** Likelihood analysis of pairwise interaction probabilities. Results are shown  
 427 for three abundance measures (see Methods): (a) density of individuals, (b) density of  
 428 flowers and (c) density of pollen. Each panel shows the dAIC values corresponding to each  
 429 of the seven probability matrices resulting from all possible combinations of abundance  
 430 (**N**), temporal overlap (**T**) and spatial overlap (**S**). The dAIC value for a null probability  
 431 matrix with homogenous interaction probability across all pairwise interactions (Null)  
 432 and the observed interaction matrix fitted to itself (**Y**) are also shown for comparison.



Figure 1:

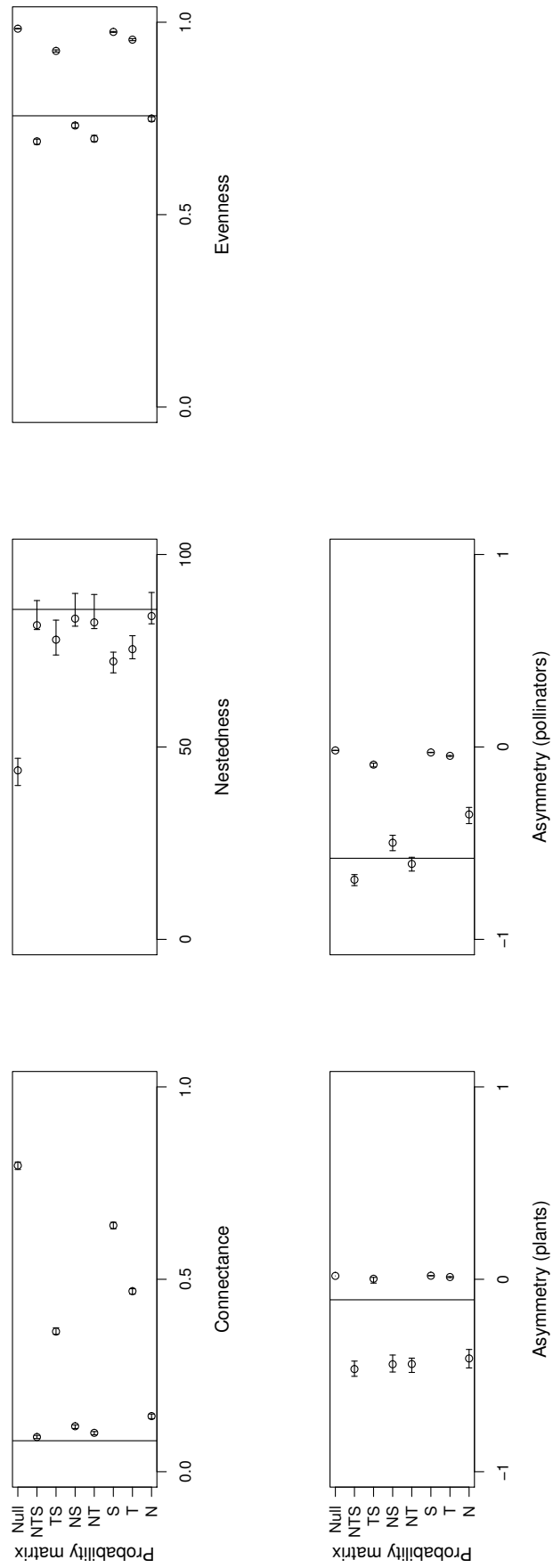


Figure 2:

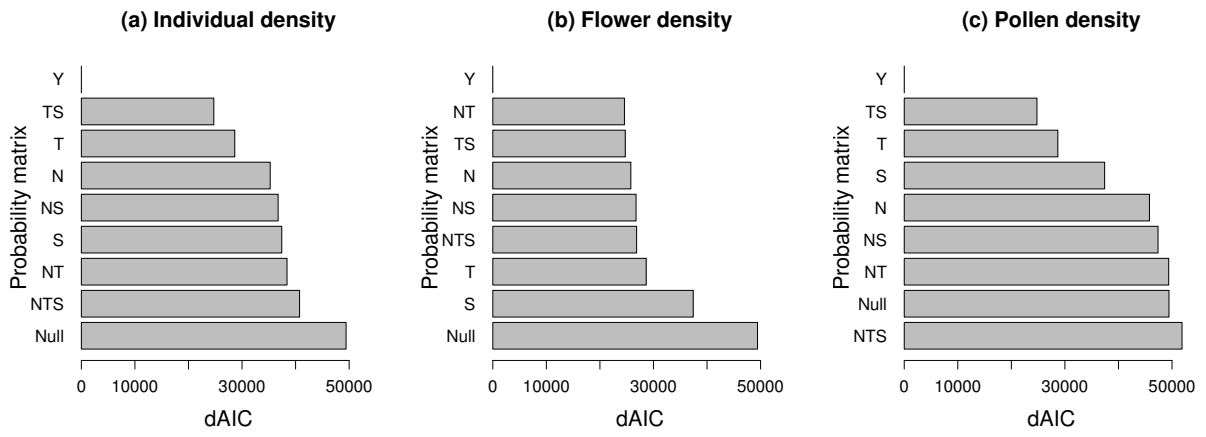


Figure 3:

## Electronic Appendices

### Appendix A. Phylogenetic analyses.

**Construction of plant and insect phylogenies.** We used the plant phylogenetic database Phylomatic (<http://www.phylodiversity.net/phylomatic>) to construct the plant phylogeny. Phylomatic is based on the Angiosperm supertree built by Davies et al. (2004), and allows inputting a list of plant species with their family affiliation to obtain a phylogenetic tree. We selected the “conservative seed plant tree” option, which leaves nodes with less of 80% support as soft polytomies. We also chose to estimate pseudo branch lengths, which Phylomatic does by calibration with dates from Wikstrom et al. (2001). The phylogeny returned by Phylomatic (Fig. A1a) was then used to calculate the phylogenetic variance–covariance matrix  $\mathbf{U}$  (Cunningham et al., 1998; Garland and Ives, 2000), using the `vcv.phylo` function in the `ape` package of R (Paradis et al., 2004).

For the insects, we assembled a family-level phylogeny from the Tree of Life (Maddison and Schulz, 2007). For families with more than two species that could be identified to genus or species, we inserted within-family phylogenies into the family-level phylogeny. Within-family phylogenies were drawn from the Tree of Life and from selected sources: Syrphidae, Ståhls et al. (2003), Mengual et al. (2008); Halictidae, Danforth et al. (2008); Megachilidae, Danforth et al. (2006). As for the plants, this phylogeny (Fig. A1b) was used to calculate the variance–covariance matrix  $\mathbf{V}$ .

**Estimation of phylogenetic signal.** We evaluated the strength of the phylogenetic signal of the two phylogenies on the interaction matrices with the methods developed by Ives and Godfray (2006), which use a linear model approach to fit the phylogenetic variance–covariance matrix to the interaction matrix. Ives and Godfray (2006) propose that the matrix of interaction strengths can be described by a linear model,  $\mathbf{Y} = b_0 + \epsilon$ , where  $\mathbf{Y}$  is the interaction matrix in vectorized form (i.e., the result of “stacking” columns into a single column vector),  $b_0$  is the phylogenetically corrected mean of interaction strength, and  $\epsilon$  is a column vector of the same dimension as  $\mathbf{Y}$ . Vector  $\epsilon$  has an associated variance–covariance matrix  $E[\epsilon\epsilon'] = \mathbf{W}$ , which is a function of the phylogenies of plants and insects. Specifically,  $\mathbf{W} = \mathbf{U} \otimes \mathbf{V}$ , where  $\mathbf{U}$  and  $\mathbf{V}$  are the phylogenetic variance–covariance matrices for plants and insects, respectively, and  $\otimes$  is the Kronecker product.

The overall strength of the phylogenetic signal can be assessed by comparing the mean squared error (MSE) calculated by fitting the model with the actual data ( $\text{MSE}_d$ ), the MSE derived under the assumption of no phylogenetic covariances (a “star” phylogeny;  $\text{MSE}_{\text{star}}$ ), and the MSE derived assuming maximum phylogenetic signal (i.e., Brownian motion evolution,  $\text{MSE}_b$ ). Lower values of MSE indicate better fit of the specific model to the data (i.e., the model with the lowest MSE leaves the smallest unexplained variance). In addition, the independent signals of the plant and the animal phylogenies,  $d_p$  and  $d_a$ , can be estimated from  $b_0$  in the linear model. A value of  $d = 0$  means no phylogenetic signal,  $d = 1$  means maximum phylogenetic signal under the assumption of Brownian motion evolution, and  $d > 1$  corresponds to disruptive selection. We used the `pblm` function in the `picante` package of R (Kembel et al., 2008) to calculate these measures of phylogenetic signal.

## References

- Cunningham, C. W., K. E. Omland, and T. H. Oakley. 1998. Reconstructing ancestral character states: a critical reappraisal. *Trends in Ecology & Evolution* **13**:361–366.
- Danforth, B., C. Eardley, L. Packer, K. Walker, A. Pauly, and F. Randrianambinintsoa. 2008. Phylogeny of Halictidae with an emphasis on endemic African Halictinae. *Apidologie* **39**:86–101.
- Danforth, B. N., S. Sipes, J. Fang, and S. G. Brady. 2006. The history of early bee diversification based on five genes plus morphology. *Proceedings of the National Academy of Sciences USA* **103**:15118–15123.
- Davies, T. J., T. G. Barraclough, M. W. Chase, P. S. Soltis, D. E. Soltis, and V. Savolainen. 2004. Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proceedings of the National Academy of Sciences USA* **101**:1904–1909.
- Garland, T. and A. R. Ives. 2000. Using the past to predict the present: confidence intervals for regression equations in phylogenetic comparative methods. *American Naturalist* **155**:346–364.
- Ives, A. R. and H. C. J. Godfray. 2006. Phylogenetic analysis of trophic associations. *American Naturalist* **168**:E1–E14.
- Kembel, S., D. Ackerly, S. Blomberg, P. Cowan, M. Helmus, and C. Webb, 2008. *picante*: Tools for integrating phylogenies and ecology. R package version 0.3-0.
- Maddison, D. R. and K.-S. Schulz, editors. 2007. The Tree of Life Web Project. Internet address: <http://tolweb.org>.
- Mengual, X., G. Ståhls, and S. Rojo. 2008. First phylogeny of predatory flower flies (Diptera, Syrphidae, Syrphinae) using mitochondrial COI and nuclear 28S rRNA genes: conflict and congruence with the current tribal classification. *Cladistics* **24**:543–562.
- Paradis, E., J. Claude, and K. Strimmer, 2004. *APE*: analyses of phylogenetics and evolution in R language.
- Ståhls, G., H. Hippa, G. Rotheray, J. Muona, and F. Gilbert. 2003. Phylogeny of Syrphidae (Diptera) inferred from combined analysis of molecular and morphological characters. *Systematic Entomology* **28**:433–450.
- Wikstrom, N., V. Savolainen, and M. W. Chase. 2001. Evolution of angiosperms: Calibrating the family tree. *Proceedings of the Royal Society B* **268**:2211–2220.

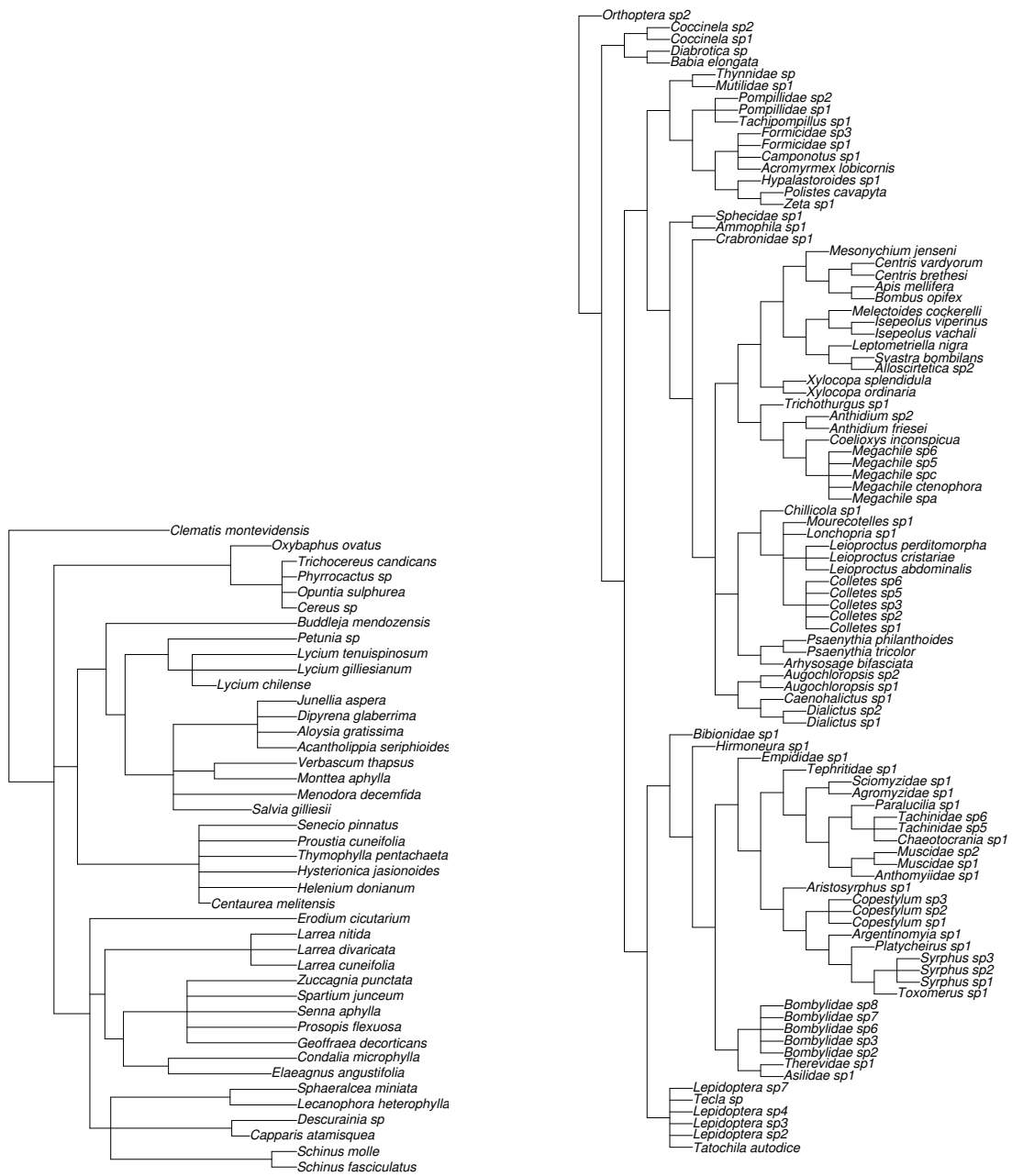


Figure A1. Phylogenies of plants and insects included in the plant–pollinator interaction network of Villavicencio Nature Reserve.

**Supplement.** R functions used for analyses.

**Main functions.** This supplement contains R functions used for the analyses presented in the main text. Main functions are `plotmat`, which draws plots of the interaction matrices (Fig. 1); `netstats`, which calculates aggregate network statistics (Fig. 2); and `mlik`, which calculates multinomial likelihood and AIC (Fig. 3). Other functions are either called by these main functions to perform operations, or do other related operations. An example of the use of these functions is available in the `bipartite` package of R (Dormann et al., 2008) (type `?vazquez.example` in the R console after installing and uploading the `bipartite` package).

## Function descriptions

`confint` Confidence intervals of a vector or matrix of simulated values. Used by `netstats` to calculate 95% confidence intervals of simulated aggregate network statistics.

`intasymm` Interaction strength asymmetry, calculated following Vázquez et al. (2007).

`intereven` Interaction evenness, calculated following Tylianakis et al. (2007).

`mgen` Matrix generating algorithm used by `netstats` to generate simulated interaction matrices according to a given probability matrix.

`mlik` Calculation of multinomial likelihood and AIC according to a given probability matrix. Usage: `mlik(imatr, m.p, par)`, where `imatr` is the observed interaction matrix, `m.p` is the probability matrix, and `par` is the number of parameters used to calculate AIC.

`netstats` Aggregate network statistics: connectance, nestedness, interaction evenness and mean interaction strength asymmetry for each group in the network. Usage: `netstats(imatr, randomize=TRUE, iter=1000, pmatr=NULL)`, where `imatr` is the observed interaction matrix and `pmatr` is the probability matrix used for generating predicted matrices. The R package `bipartite` (Dormann et al., 2008) is required by `netstats` to calculate nestedness.

`plotmat` Graphic function to plot interaction matrices as in Fig. 1. Usage: `plotmat(imatr, xlabel="PLANTS", ylabel="POLLINATORS", cexmin=0.2, cexmax=2, sortm=TRUE, sqrt=TRUE)`, where `imatr` is the interaction matrix to be plotted, `xlabel` and `ylabel` are the texts for the  $x$  and  $y$  axes, `cexmin` and `cexmax` are the minimum and maximum sizes of circles representing pairwise interactions, `sortm` indicates whether matrix should be sorted (`sortm=TRUE`) or left in the original order (`sortm=FALSE`), and `sqrt` indicates whether interaction frequencies (matrix cells) are to be plotted in untransformed or square-root transformed.

`quant2bin` Transformation of quantitative matrix into binary. Used by `netstats` to calculate connectance.

`sortmatr` Matrix sorting algorithm used by `plotmat`. The matrix is sorted according to row and column totals, so that nestedness can be visualized.

`sortmatrext` Matrix sorting algorithm used as an alternative to `sortmatr`. Here, the matrix is sorted according to external vectors instead of row and column totals.

## References

- Dormann, C. F., B. Gruber, and J. Fründ. 2008. Introducing the bipartite package: analysing ecological networks. *R News* **8/2**:8–11.
- Tylianakis, J. M., T. Tscharntke, and O. T. Lewis. 2007. Habitat modification alters the structure of tropical host-parasitoid food webs. *Nature* **445**:202–205.
- Vázquez, D. P., C. J. Melián, N. M. Williams, N. Blüthgen, B. R. Krasnov, and R. Poulin. 2007. Species abundance and asymmetric interaction strength in ecological networks. *Oikos* **116**:1120–1127.