

Degree distribution in plant–animal mutualistic networks: forbidden links or random interactions?

Diego P. Vázquez, National Center for Ecological Analysis and Synthesis, Univ. of California, Santa Barbara, 735 State Street, Suite 300, Santa Barbara, CA 93101, USA (vazquez@nceas.ucsb.edu).

Recent studies show that ecological interaction networks depart from the “scale-free” topologies observed in many other real world networks. Such a departure has been hypothesized to result from non-matching biological attributes of species, such as phenology or morphology, that prevent the occurrence of certain interactions (“forbidden links”). Here I compare the topology of 17 plant–animal mutualistic networks with that predicted by a simple null model that assumes that a species’ degree (number of interspecific interactions) is a function of its frequency of interaction. The topology predicted by this null model is strikingly close to that observed in the real networks. Thus, this null model provides a simple alternative interpretation of patterns observed in ecological interaction networks that does not require the existence of non-matching species traits.

Ecologists often have depicted interactions among coexisting species as a network, in which nodes represent species and links between pairs of nodes represent interspecific interactions (Paine 1966, May 1974, Pimm 1982, Jordano 1987, Polis 1991). A primary goal of this research has been the search for regularities in network topology and the mechanisms behind such regularities (Cohen and Newman 1985, Jordano 1987, Martinez 1992, Williams and Martinez 2000). Two recent studies (Dunne et al. 2002, Jordano et al. 2003) have shown that ecological interaction networks depart from the scale-free topology observed in many other real world networks (such as the world wide web, the internet, networks of scientific citations and collaborators, and metabolic and protein networks; Barabási and Albert 1999, Albert and Barabási 2002, Newman 2003).

Under a scale free topology, the number of links per node (hereafter “degree”) follows a powerlaw distribution, usually represented as $P(k) \sim k^{-\gamma}$, where k is the degree, $P(k)$ is the cumulative degree distribution (the proportion of nodes with k or more links), and γ is a constant (Amaral et al. 2000, Newman 2003). Such a distribution may result from a process of network growth through preferential attachment, whereby new nodes are added to the network by preferentially

attaching to highly connected nodes (Barabási and Albert 1999). In contrast, degree distribution in ecological interaction networks typically decays faster than expected under a power law (Dunne et al. 2002, Jordano et al. 2003). In the particular case of plant–animal mutualistic interactions, most networks follow a “broad scale” distribution (Jordano et al. 2003), best described by a power law distribution with an exponential decay, $P(k) \sim k^{-\gamma} \exp(-k/k_x)$, where k_x is a constant and k , $P(k)$ and γ are defined as above.

Jordano et al. (2003) suggested the existence of “forbidden links” as the underlying causes of departure from the scale free degree distribution. These forbidden links are presumably the result of biological attributes of species that prevent the occurrence of some interactions. For example, interactions would not be possible between species with non-overlapping phenologies or non-matching morphologies. Jordano et al. suggest that this structural constraint would impose an upper limit on the number of links per node, causing the degree distribution to drop off more rapidly than under a power law.

Although network growth through preferential attachment and the existence of forbidden links are certainly possible, it is unclear whether these mechanisms are sufficient explanations of observed degree distributions in plant–animal mutualistic networks. First, although forbidden links likely contribute to generate the observed number of links in a given network (Jordano 1987), interactions among temporally overlapping species seem to be “allowed” in many cases (Waser et al. 1996, Richardson et al. 2000). Second, and more importantly, forbidden links resulting from phenological or morphological constraints are equally likely to affect any species, not just the most connected ones, and it is unclear whether this assembly constraint would necessarily lead to a decay in the tail of the degree distribution.

In this article I propose a different, more parsimonious mechanism to explain the observed degree distribution of plant–animal mutualistic networks. I use a previously published null model of plant–animal mutualistic interactions that assumes that the degree of a species is a function of its frequency of interaction (i.e. the number of times a species was recorded interacting in the field). Therefore, the model assumes that species traits are irrelevant in determining inter-specific interactions directly, and it thus provides a simple explanation against which other models could be compared.

The data

I studied degree distribution in twelve plant–pollinator and five plant–frugivore interaction datasets from published community-wide studies (Appendix 1). Data come from field studies at local communities in different parts of the world, in which the interactions among most plants at a particular locality and their animal mutualists (pollinators or frugivores) were recorded. Data are available through the Interaction Web Database (<http://www.nceas.ucsb.edu/interactionweb>).

For each dataset, I constructed a binary interaction matrix, in which rows and columns represent animal and plant species, respectively, a cell with a “1” represents an interspecific interaction (hereafter “link”) between an animal and a plant species, and a cell with a “0” represents no interspecific interaction. In addition, the datasets contained information that allowed estimation of the frequency with which each species was observed to engage in an individual interaction (“frequency of interaction;” i.e. the total number of visits of any pollinator species received by a plant species, or the total number of visits made by a particular pollinator species to any plant species).

The null model

I used a previously developed randomization model (i.e. the “null model 2” of Vázquez and Aizen 2003, 2004) to simulate interactions among plants and their animal mutualists. This model assumes that the network is static (i.e. the number of nodes does not grow with time), and requires that all species interact with at least one other species and that the total number of links in the randomized networks is the same as originally observed. The model incorporates a correlation between a species’ frequency of interaction and its degree that underlies most data (Appendix 1). To this end, the model assumes that the probability of a species to be linked to other species is roughly proportional to its relative frequency of interaction. Thus, species with a high frequency of

interaction should have a higher degree than species with a lower frequency of interaction (Vázquez and Aizen 2003, 2004).

The null model, implemented in Matlab (MathWorks 1999), is based on the randomization of the binary interaction matrix according to species-specific probabilities. To ensure that each species has at least one link, the algorithm first assigns one link to each species by drawing one species from the pool with probability equal to the relative frequency of interaction of each potential interaction partner. Subsequent links are then assigned by selecting one plant and one pollinator species according to their relative frequencies of interaction; this process is repeated until the number of links originally observed is reached. Thus, the probability of interaction between an animal species i and a plant species j is approximately $p_{ij} = p_i p_j$, where p_i and p_j are the relative interaction frequencies of animal i and plant species j (see Vázquez and Aizen [2003, 2004] for more details about the algorithm). Randomizations were repeated 1000 times for each dataset. The predicted cumulative degree distribution was calculated as the mean of the 1000 randomized degree distributions.

Assessment of model fit

Following previous studies (Amaral et al. 2000, Dunne et al. 2002, Jordano et al. 2003), I calculated the cumulative degree distribution for each original dataset excluding categories that had zero frequency in the noncumulative probability distribution histogram, and then assessed the fit of the degree distribution predicted by the null model. Although the above procedure can be problematic (because of non-independence of data points and inflation of model fit; Newman 2003), doing so makes my results directly comparable with those of previous studies (Dunne et al. 2002, Jordano et al. 2003).

One important point about the approach taken here is that it does not involve any parameter-fitting procedure. Thus, although the model does have several parameters (the probabilities associated to each species, the observed number of links), the values of these parameters are set a priori; i.e. I specify the model before comparing it to the data. This procedure makes the test of the model particularly strong (McGill 2003).

In some cases, the categories with the highest degree had a predicted cumulative probability equal to zero. The finite number of iterations used in the randomization resulted in some cases in zero cumulative probabilities for the highest degree categories, which made logarithmic transformation impossible. However, in the limit, when the number of iterations tends to infinity, all categories should have a nonzero probability. Therefore, I assumed that the lowest possible associated

cumulative probability was 1 over the number of iterations (i.e. 10^{-3}).

Percent variance in the data explained by the model prediction was estimated as $R^2 = \left(\frac{\sum Y_i \hat{Y}_i}{\sqrt{\sum Y_i^2 \sum \hat{Y}_i^2}} \right)^2$, where Y_i and \hat{Y}_i are the observed and predicted values of the cumulative degree distribution for each bin i of the cumulative degree histogram. (Note that in traditional least square regression, such metric is provided by the

coefficient of multiple determination, R^2 , calculated as $R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}$, where Y_i and \hat{Y}_i are defined as above and \bar{Y} is the mean of the observed and predicted values, which are assumed to be the same as a result of the fitting procedure. However, because my modeling approach did not involve fitting of any free parameters, the means of observed and predicted distributions are not necessarily the same, which makes the use of this form of R^2 inappropriate.)

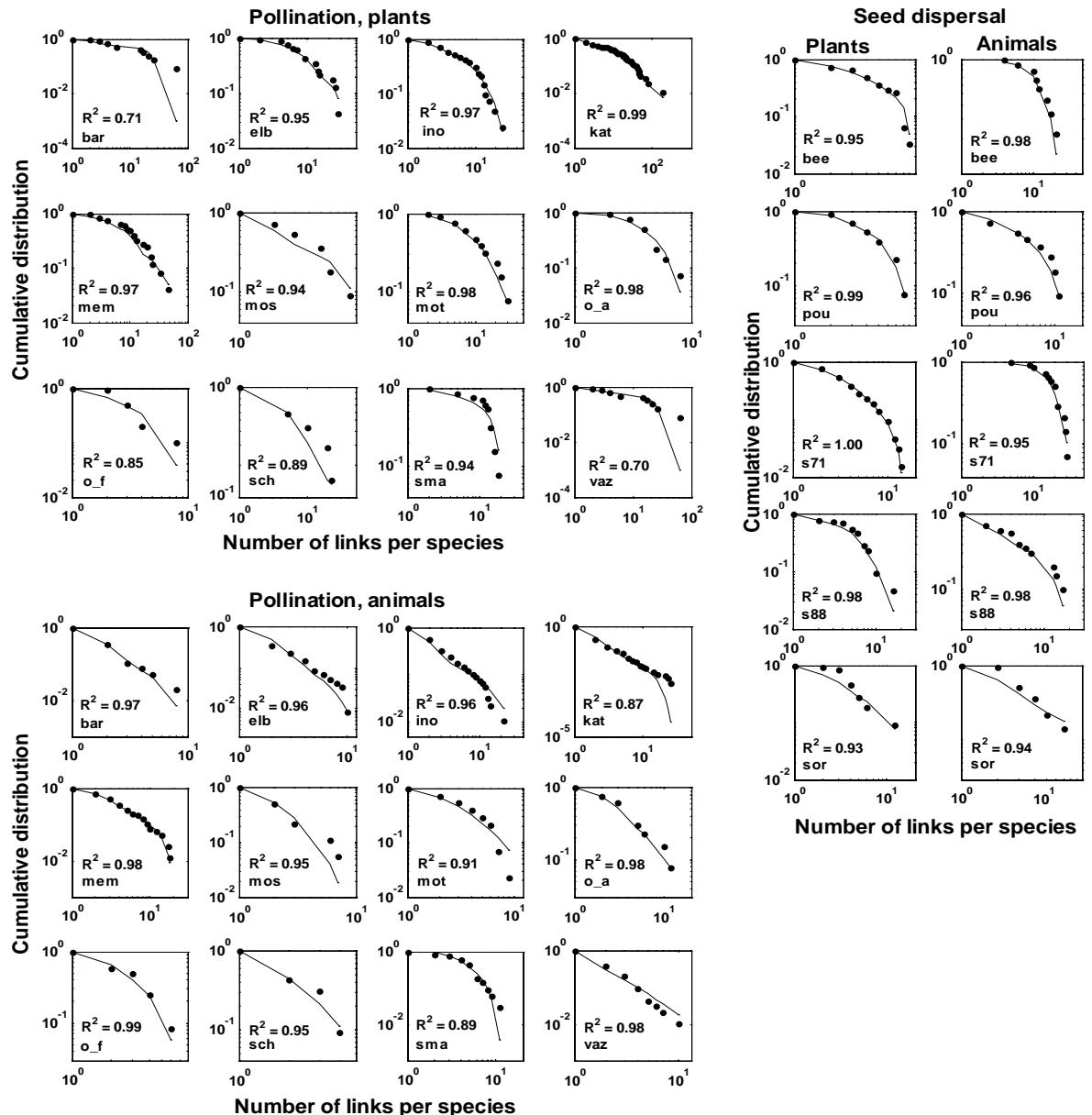


Fig. 1. Degree distribution in 12 plant–pollinator and 5 plant–seed disperser bipartite interaction networks. Dots represent observed relative frequencies of species with k or more links; lines are expected degree distributions predicted by the randomization model. Dataset codes are indicated in the lower-left corner of each panel (see Supplementary Information for full references). The R^2 statistic, estimating the proportion of variance in the data explained by the model, is also given for each dataset.

Results and discussion

Results of my analysis suggest that the null model provides an accurate prediction of cumulative degree distribution. Degree distribution predicted by the null model is very close to the observed distribution in most datasets (Fig. 1). Thus, the existence of species traits that act to prevent some pairs of species from interacting (i.e. forbidden links) is not necessary to explain the observed topology of real ecological networks. Rather, these results demonstrate that the broad scale patterns observed both in plant–animal mutualistic networks and in food webs (Dunne et al. 2002, Jordano et al. 2003) can be explained by an underlying correlation between frequency of interaction and degree.

Of course, showing that a given model fits the data well is by no means a demonstration that the mechanism actually operates; more than one mechanism can produce any given pattern (McGill 2003). However, because of the simplicity of the model and its close fit to the data, this mechanism should be considered at least as good an explanation of observed patterns as more sophisticated mechanisms proposed in previous studies, including network growth through preferential attachment and the existence of forbidden interactions (Jordano et al. 2003). Testing for the effect of forbidden links may be difficult, and much more detailed information about the biology of species interactions will be necessary; until these data are available, the hypothesis that topological patterns of interaction networks are simply the result of a stochastic process like the one assumed in the present null model cannot be rejected.

Several important implications emerge from this study. First, my results should warn against the over-interpretation of topological patterns observed in ecological interaction networks. The fact that the assumption that frequently interacting species have a higher degree than rarely interacting species is sufficient to explain observed topological patterns suggests that most phenotypic characteristics of interacting species may be irrelevant in determining broad patterns of interspecific interactions. However, species frequency of interaction does probably emerge from species traits, such as those determining species abundance, flower and fruit attractiveness to pollinators and frugivores, and animal mobility (Vázquez and Aizen 2005). Thus, species traits may be important in structuring interaction networks, but in a fundamentally different way from the one envisioned in previous studies (Bascompte et al. 2003, Jordano et al. 2003).

Second, the fact that this simple null model consistently generates patterns observed in real communities suggests that we may be looking at the wrong patterns. The aggregated statistical descriptions that have been the

focus of recent work on networks (including my own work) remove much of the details of the interactions; thus, that these simplified patterns are well explained by equally simplified mechanisms does not mean that more complex patterns that require more complex explanations do not exist. I suggest that a better understanding of ecological interactions will come from studies that attempt to delve deeper into the details of interactions. Particularly, as many have previously argued (Cohen et al. 1993, Memmott 1999, Borer et al. 2002), it will be important to move beyond binary representations of interactions to quantitative measures involving interaction strength.

Acknowledgements – I thank Luis Amaral, Karl Cottenie, Jennifer Dunne, and Eric Torbet for discussion, and Elizabeth Borer, Bill Langford, Helene Muller-Landau, Pedro Jordano and Nick Waser for useful comments on the manuscript. Jeremy Goldberg, Rupesh Naik and the Interaction Web Database facilitated access to the data used in this study. I was supported by a postdoctoral fellowship at the National Center for Ecological Analysis and Synthesis (a center funded by NSF Grant #DEB-0072909, the University of California, and the Santa Barbara campus).

References

- Albert, R. and Barabási, A. L. 2002. Statistical mechanics of complex networks. – *Rev. Modern Physics* 74: 47–97.
- Amaral, L. A. N., Scala, A., Barthelemy, M. et al. 2000. Clases of small-world networks. – *Proc. Natl Acad. Sci. USA* 97: 11149–11152.
- Barabási, A. L. and Albert, R. 1999. Emergence of scaling in random networks. – *Science* 286: 509–512.
- Bascompte, J., Jordano, P., Melián, C. J. et al. 2003. The nested assembly of plant–animal mutualistic networks. – *Proc. Natl Acad. Sci. USA* 100: 9383–9387.
- Borer, E. T., Anderson, K., Blanchette, C. A. et al. 2002. Topological approaches to food web analyses: a few modifications may improve our insights. – *Oikos* 99: 397–401.
- Cohen, J. E. and Newman, C. M. 1985. A stochastic theory of community food webs. I. Models and aggregated data. – *Proc. R. Soc. Lond. B* 224: 421–448.
- Cohen, J. E., Beaver, R. A., Cousins, S. H. et al. 1993. Improving food webs. – *Ecology* 74: 252–258.
- Dunne, J. A., Williams, R. J. and Martinez, N. D. 2002. Food-web structure and network theory: the role of connectance and size. – *Proc. Natl Acad. Sci. USA* 99: 12917–12922.
- Jordano, P. 1987. Patterns of mutualistic interactions in pollination and seed dispersal: connectance, dependence asymmetries, and coevolution. – *Am. Nat.* 129: 657–677.
- Jordano, P., Bascompte, J. and Olesen, J. M. 2003. Invariant properties in coevolutionary networks of plant–animal interactions. – *Ecol. Lett.* 6: 69–81.
- Martinez, N. D. 1992. Constant connectance in community food webs. – *Am. Nat.* 139: 1208–1218.
- MathWorks. 1999. Matlab. – The MathWorks, Inc.
- May, R. M. 1974. Stability and complexity in model ecosystems. – Princeton Univ. Press.
- McGill, B. 2003. Strong and weak tests of macroecological theory. – *Oikos* 102: 679–685.
- Memmott, J. 1999. The structure of a plant–pollinator food web. – *Ecol. Lett.* 2: 276–280.

- Newman, M. E. J. 2003. The structure and function of complex networks. – *Siam Rev.* 45: 167–256.
- Paine, R. T. 1966. Food web complexity and species diversity. – *Am. Nat.* 100: 65–75.
- Pimm, S. L. 1982. Food webs. – Chapman and Hall.
- Polis, G. A. 1991. Complex trophic interactions in deserts: an empirical critique of food-web theory. – *Am. Nat.* 138: 123–155.
- Richardson, D. M., Allsopp, N., D’Antonio, C. M. et al. 2000. Plant invasions – the role of mutualisms. – *Biol. Rev.* 75: 65–93.
- Vázquez, D. P. and Aizen, M. A. 2003. Null model analyses of specialization in plant–pollinator interactions. – *Ecology* 84: 2493–2501.
- Vázquez, D. P. and Aizen, M. A. 2004. Asymmetric specialization: a pervasive feature of plant–pollinator interactions. – *Ecology* 85: 1251–1257.
- Vázquez, D. P. and Aizen, M. A. 2005. Community-wide patterns of specialization in plant–pollinator interactions revealed by null-models. – In: Waser, N. M. and Ollerton, J. (eds), *Specialization and generalization in plant–pollinator interactions*. Univ. of Chicago Press, in press.
- Waser, N. M., Chittka, L., Price, M. V. et al. 1996. Generalization in pollination systems, and why it matters. – *Ecology* 77: 1043–1060.
- Williams, R. J. and Martinez, N. D. 2000. Simple rules yield complex food webs. – *Nature* 404: 180–183.

Appendix 1. Datasets included in the study. Data are available through the Interaction Web Database (<http://www.nceas.ucsb.edu/interactionweb>)

Interaction type	Dataset code	Number of species		f–s correlation [§]		Reference
		Plants	Animals	Plants	Animals	
Pollination	bar	12	102	0.78	0.78	Barrett and Helenurm (1987)
	elb	23	118	0.97	0.90	Elberling and Olesen (1999)
	ino	42	91	0.72	0.85	Inouye and Pyke (1988)
	kat	89	679	0.96	0.87	Kato et al. (1990)
	mem	25	79	0.95	0.91	Memcott (1999)
	mos	11	18	0.90	0.70	Mosquin and Martin (1967)
	mot	13	44	0.92	0.85	Motten (1982)
	o_a	14	13	0.08	0.57	Olesen et al. (2002)
	o_f	10	12	0.85	0.91	Olesen et al. (2002)
	sch	7	32	0.88	0.89	Schemske et al. (1978)
	sma	13	34	0.88	0.31	Small (1976)
	vaz	14	93	0.72	0.73	Vázquez and Simberloff (2002, 2003) [†]
Seed dispersal	bee	31	9	0.89	0.96	Beehler (1983)
	pou	13	11	0.86	0.98	Poulin et al. (1999)
	s71	65	14	0.86	0.89	Snow and Snow (1971)
	s88	22	20	0.75	0.91	Snow and Snow (1988)
	sor	12	14	0.70	0.38	Sorensen (1981)

[§]Pearson’s correlation coefficient between a species’ frequency of interaction (f) and the number of species of interaction partners (s), calculated separately for plants and animals. Correlation coefficients with associated probability $P < 0.05$ are highlighted in bold.

[†]This dataset consisted of data from eight sites with the same community type within a 25 km radius; because analyses done separately for each site and for all sites pooled gave similar results, I report results for the pooled data only.

References for Appendix 1

- Barrett, S. C. H. and Helenurm, K. 1987. The reproductive biology of boreal forest herbs. I. Breeding systems and pollination. – *Can. J. Bot.* 65: 2036–2046.
- Beehler, B. 1983. Frugivory and polygamy in birds of paradise. – *Auk* 100: 1–12.
- Elberling, H. and Olesen, J. M. 1999. The structure of a high latitude plant–flower visitor system: the dominance of flies. – *Ecography* 22: 314–323.
- Inouye, D. W. and Pyke, G. H. 1988. Pollination biology in the Snowy Mountains of Australia: comparisons with montane Colorado, USA. – *Aust. J. Ecol.* 13: 191–210.
- Kato, M., Makutani, T., Inoue, T. et al. 1990. Insect–flower relationship in the primary beech forest of Ashu, Kyoto: an overview of the flowering phenology and seasonal pattern of insect visits. – *Contr. Biol. Lab. Kyoto Univ.* 27: 309–375.
- Memcott, J. 1999. The structure of a plant–pollinator food web. – *Ecol. Lett.* 2: 276–280.
- Mosquin, T. and Martin, J. E. H. 1967. Observations on the pollination biology of plants on Melville Island, N.W.T., Canada. – *Can. Field Nat.* 81: 201–205.
- Motten, A. F. 1982. Pollination ecology of the spring wildflower community in the deciduous forests of piedmont North Carolina. Dept of Botany. – Duke Univ., p. xiii + 205

- Olesen, J. M., Eskildsen, L. I. and Venkatasamy, S. 2002. Invasion of pollination networks on oceanic islands: importance of invader complexes and endemic super generalists. – *Div. Distr.* 8: 181–192.
- Poulin, B., Wright, S. J., Lefebvre, G. et al. 1999. Interspecific synchrony and asynchrony in the fruiting phenologies of congeneric bird-dispersed plants in Panama. – *J. Trop. Ecol.* 15: 213–227.
- Schemske, D. W., Willson, M. F., Melampy, M. N. et al. 1978. Flowering ecology of some spring woodland herbs. – *Ecology* 59: 351–366.
- Small, E. 1976. Insect pollinators of the Mer Bleue peat bog of Ottawa. – *Can. Field Nat.* 90: 22–28.
- Snow, B. K. and Snow, D. W. 1971. The ecology of tanagers and honeycreepers in Trinidad. – *Auk* 88: 291–322.
- Snow, B. K. and Snow, D. W. 1988. *Birds and berries.* – Carlton.
- Sorensen, A. E. 1981. Interactions between birds and fruit in a temperate woodland. – *Oecologia* 50: 242–249.
- Vázquez, D. P. and Simberloff, D. 2002. Ecological specialization and susceptibility to disturbance: conjectures and refutations. – *Am. Nat.* 159: 606–623.
- Vázquez, D. P. and Simberloff, D. 2003. Changes in interaction biodiversity induced by an introduced ungulate. – *Ecol. Lett.* 6: 1077–1083.