The modularity of pollination networks

Jens M. Olesen*†, Jordi Bascompte‡, Yoko L. Dupont*, and Pedro Jordano‡

*Department of Biological Sciences, University of Aarhus, Ny Munkegade Block 1540, DK-8000 Aarhus C, Denmark; and ‡Integrative Ecology Group, Estación Biológica de Doñana, Consejo Superior de Investigaciones Científicas, Pabellón del Perú, Avenida María Luisa s/n, E-41013 Seville, Spain

In natural communities, species and their interactions are often organized as nonrandom networks, showing distinct and repeated complex patterns. A prevalent, but poorly explored pattern is ecological modularity, with weakly interlinked subsets of species (modules), which, however, internally consist of strongly connected species. The importance of modularity has been discussed for a long time, but no consensus on its prevalence in ecological networks has yet been reached. Progress is hampered by inadequate methods and a lack of large datasets. We analyzed 51 pollination networks including almost 10,000 species and 20,000 links and tested for modularity by using a recently developed simulated annealing algorithm. All networks with >150 plant and pollinator species were modular, whereas networks with <50 species were never modular. Both module number and size increased with species number. Each module includes one or a few species groups with convergent trait sets that may be considered as coevolutionary units. Species played different roles with respect to modularity. However, only 15% of all species were structurally important to their network. They were either hubs (i.e., highly linked species within their own module), connectors linking different modules, or both. If these key species go extinct, modules and networks may break apart and initiate cascades of extinction. Thus, species serving as hubs and connectors should receive high conservation priorities.

coevolution | compartment | module | nestedness | species role

Biodiversity encompasses not just species but also interactions among species. Within habitats, species and their interactions assemble into large, complex ecological networks. Such networks are rich in structural heterogeneity (1). Understanding network structure and its underlying causes are essential parts of any study of biodiversity and its responses to disturbances, yet it is a conceptual and methodological challenge to address these problems in highly diversified communities with thousands of interactions.

Moving through an ecological network of species and their connecting links, one traverses a heterogeneous universe of link-dense and -sparse areas. Link-dense regions are termed compartments (2) or, here, modules (3), whereas link-sparse regions demarcate their boundaries. Species within a module are linked more tightly together than they are to species in other modules. The extent to which species interactions are organized into modules is termed the modularity of the network. Modularity may reflect habitat heterogeneity, divergent selection regimes, and phylogenetic clustering of closely related species (4, 5), leading to nonrandom patterns of interaction and ultimately contributing to the complexity of ecological networks. Modules with their tightly linked species may even be the long-sought key units of coevolution, in which reciprocal selection leads to trait convergence in unrelated species (6). However, modularity has been notoriously difficult to demonstrate either because of its rarity or because of a lack of sufficiently strong module-detecting algorithms (1).

In ecology, Pimm and Lawton (4) presaged recent methodological developments by constructing a modularity statistic, which they compared with a distribution of randomly generated webs. Later, Raffaelli and Hall (7) used the distribution of a trophic similarity index between species pairs to look for mod-

ules in an estuary web. Dicks *et al*. (8), adopting the latter method (7), analyzed two pollination networks and found that they were modular. Today, strong modularity-detecting algorithms are used outside ecology, e.g., in social sciences (9), information theory (10), and network studies of airlines and biochemistry (11–13). Recently, one of these algorithms (14) was used to detect modularity in five food webs (15). We used an algorithm, here termed SA (11–13), which is based on simulated annealing [see supporting information (SI) *Materials and Methods* for a review of different methods] in a search for modules in pollination networks. Together with other properties of such mutualistic networks, e.g., nestedness and asymmetry (16–20), modularity is expected to be an essential ingredient of network complexity. Thus, modularity may play a critical role in both the functioning of these networks, e.g., to their stability (21, 22) and in the potential for coevolution of plant and animal species (6). An identification of modules can tell us what highly connected groups of species look like and perhaps also at what scale coevolution is taking place. This may ultimately assist us in the maintenance of biodiversity by preserving key groups of taxa playing a fundamental role to network persistence.

A long-lasting tenet in mutualistic plant–animal interaction studies is the concept of syndromes (23, 24), i.e., that species converge on correlated suites of traits shaped by similar interaction patterns. Such a process may result in a heterogeneous web of interactions, with taxonomically or functionally related taxa packed into distinct modules. Understanding how species interact and evolve at the modular level may thus be a key to an understanding of trait evolution. Many pollination studies implicitly assume modularity by focusing on subwebs of interacting species sharing a syndrome (e.g., hummingbirds and their flowers; ref. 25). Thus, a stronger, more explicit focus on modules in mutualistic networks as study objects may open up routes to a broader study of functional diversity and coevolutionary units.

Here, we focus on the modularity of entire pollination networks, i.e., all interacting species within an area are included. Our aims are to (*i*) to estimate the level of modularity in pollination networks, (*ii*) to identify different types of modules in networks, and (*iii*) to classify species into different functional roles with respect to their position within and among modules.

Results

For each pollination network in the dataset, we ran the modularity-detecting algorithm SA (see *Materials and Methods*). It produced a modularity index *M*, which was a measure of the degree to which the network was organized into clearly delimited modules. The algorithm also provided (*i*) the significance level of *M* of the real network by comparing its value to that of similar-sized random networks, (*ii*) the number of modules per
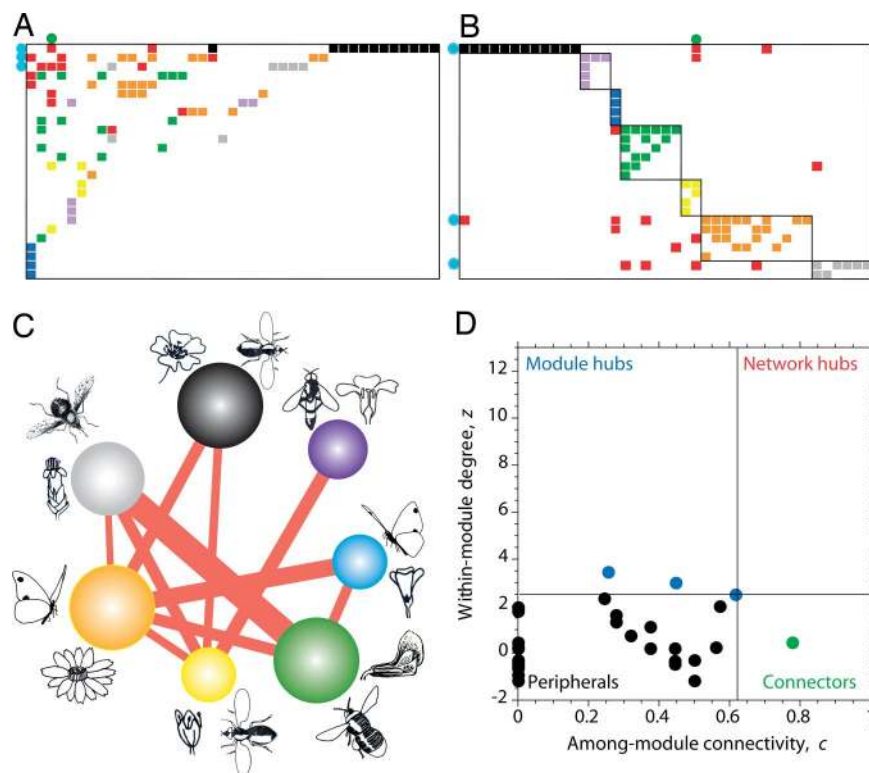
Fig. 1. Pollination network structure and species roles. The example is from the Andes (28). (*A*) Nested matrix version of the network, with plant species in columns and pollinator species in rows sorted from the upper left corner according to descending species degree. Colored cells are links between species. (*B*) Modular matrix version of the same network as in *A*; species are now sorted according to their modular affinity (order of modules is arbitrary). Red cells are species links gluing the seven modules together into a coherent network, and nonred cells are links within modules (links of the different modules can be identified in *A* by their color). (*C*) Graph of modules. Module links are weighted by both the number of species links between modules and number of species within modules; vignettes show dominant pollinator and flower type: black module: Diptera species and mainly white flowers; purple module: small-to-medium-sized beetles, flies, an ant, and small, white/yellow flowers; blue module: butterflies and one plant species, *Oxalis* species; green module: bees, birds, and large, mainly yellow flowers with a closed morphology; yellow module: large flies and small, yellow umbellifer flowers; orange module: butterflies, a large fly, and white/yellow/ pink/purple flowers; and gray module: large flies and mainly small, white flowers. (*D*) *zc*-plot of species roles with three module hubs and one connector, but no network hub. In *A* and *B*, species acting as module hubs and connectors are shown as blue and green dots, respectively, just outside the matrix border. Notice that the connector species in *B* (green dot) has >50% of its links outside its own module.

network, and (*iii*) the content of species of each module (see *Materials and Methods*). In total, 51 networks were analyzed, and 29, or 57%, were significantly modular, i.e., they had a modularity index M significantly higher than that of random networks (see SI Table 1). Mean $\overline{M} \pm$ SD was 0.52 ± 0.071 ($n = 29$ networks). All networks of >150 species were modular and all <50 species were nonmodular. Thus, modular networks encompassed by far most of the species and links in our sample (8,233 species and 16,799 links, or 85%).

Our subsequent analysis included only the 29 significantly modular networks. These networks had, on average, 8.8 ± 3.7 modules—with a maximum of 19 modules for the Amami Islands (26) and a minimum of five for a temperate forest meadow (27) and heathland (Y.L.D., unpublished work). Modularity M was independent of network size S, i.e., total number of species of pollinators A, and plants P ($n = 29$ networks: $F_{1,27} = 0.0003, P \geq 0.99$), whereas, number of modules in a network N_M, and mean module size S_M, i.e., number of species per module, both increased with S ($n = 29$: $F_{1,27} = 9.9, P \leq 0.004$; and $F_{1,27} = 45.5, P \leq 0.0001$, respectively).

Most links in a network were among species within the same module (on average 60% of all links I in a network). Thus, mean connectance C_M, within a module (number of observed links/ number of possible links in the module, excluding links to other modules), became very high ($42 \pm 28\%$), whereas mean con-

nectance C ($=100I/(AP)$), for the total networks was much lower ($7 \pm 4\%$).

All 51 networks were also tested for nestedness, i.e., a widespread pattern where specialists interact with a subset of the species that the more generalized species interact with (16). All networks, except five small ones, were significantly nested (see SI Table 1). Levels of nestedness N and modularity M were uncorrelated [$n = 29$ networks: $F_{1,27} = 2.96, P \leq 0.10$, negative trend; in this analysis, we corrected for among-network variation in S and I by using relative nestedness $(N - \overline{N}_{random})/\overline{N}_{random}$ and relative modularity $(M - \overline{M}_{random})/\overline{M}_{random}$, where $\overline{N}_{random}$ and $\overline{M}_{random}$ were average nestedness and average modularity, respectively, of the random runs]. A nested matrix thus appeared to be built on modules assembled by interactions that connected them; largely generalist species (Fig. 1*A*) were involved in these interactions. In Fig. 1*A* and *B*, an example of a real network is shown in both its nested and modular matrix version, and in Fig. 1*C*, its modular structure is depicted as a graph. The nested pattern in Fig. 1*A* resulted from the assembly of distinct modules (nonred colors in Fig. 1*B*) "glued" together by interactions among modules (red cells).

The 29 modular networks consisted of 254 modules. Individual modules differed in size and shape because of variation in species number S_M ($= A_M + P_M$, i.e., the sum of pollinator and plant species in a module) and species ratio A_M/P_M. On average, $\pm$SD, a module contained 32 ± 34 species, namely, 26 pollinator species and six plant species, i.e., $A_M/P_M = 4.2$. Thirty-six (14%)
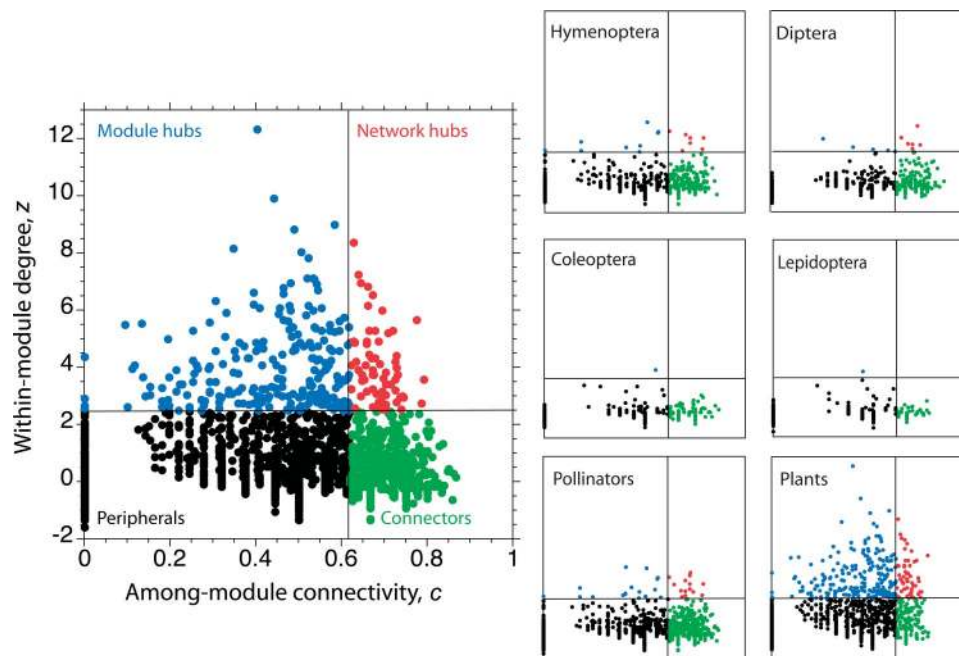
Fig. 2. Distribution of pollinator and plant species according to their network role. Each dot represents a species; large pane includes all 8,233 species from the 29 modular networks. Small panes show role distribution of selected groups of species.

of all modules were isolated species groups without any links to the remaining network. Twenty-one of these isolates were small 1:1 modules, consisting of only one pollinator species interacting with one plant species. Twenty-nine (11%) of all modules were structurally very asymmetrical or star-shaped, consisting of one generalist hub species, most often a plant species, linked to a swarm of peripheral pollinator species (range 3–51) linked only to the hub. The hub, however, had links to other modules. The remaining 189 (74%) of the modules varied a lot in size and shape (for examples of module types, see Fig. 1*B*). Individual modules might be dominated by a few species, taxonomically or functionally delimited; for example, in the subandine scrub (28), there was an 8:18-module (i.e., a module of eight pollinators interacting with 18 plants), which was dominated by three or more floral oil-collecting bees, namely, *Centris* spp. and *Tapinotaspis* species, together with five or more floral oil-producing plants, namely, *Sisyrinchium* species and *Calceolaria* species. In a Jamaican network (29), one of the seven modules was a 5:3 module consisting of four birds (Streamertail Hummingbird, Bananaquit, Orangequit, and Jamaican Blackbird), one *Anolis* lizard, and three plant species with red/purple-colored flowers.

Topologically, species played different roles in the networks. The algorithm SA described the role of each species by its within-module degree z, i.e., its standardized number of links to other species in the same module, and its among-module connectivity c, i.e., the level to which the species was linked to other modules (see *Materials and Methods*). For each species in a network, SA calculated these parameter values (11–13). Species with both a low z and a low c were peripheral species or specialists, i.e., they had only a few links and almost always only to species within their module. Species with either a high value of z or c were generalists. These included module hubs, i.e., highly connected species linked to many species within their own module (high z, low c), and connectors linking several modules (low z, high c). Species with both a high z and a high c were network hubs or super generalists, acting as both connectors and module hubs. An example of this mapping of species roles is given for a single network in Fig. 1*D*. The scatter of zc-roles of all species from all modular networks is shown in Fig. 2.

Eighty-five percent of all species were peripherals with most of their links inside their module (72% of these even had a $c = 0$, i.e., they had no links at all outside their own module) and 15% were generalists, i.e., 3% were module hubs, 11% connectors, and only 1% network hubs. The strongest module hubs were plants, like umbellifers (e.g., refs. 26, 30, and 31). Connectors were mainly beetles, flies, and small-to-medium-sized bees (e.g., refs. 32–34). Most network hub pollinators were social bees, especially *Apis* spp. (31, 35) and *Bombus* spp. (27) or large solitary bees, e.g., *Xylocopa* sp. (34) and a few Diptera species (Fig. 2).

Discussion

Above a size of 50 species, all pollination networks got significantly nested (16) and some also became modular, and passing 150 species, they were always both significantly nested and modular (see SI Table 1). Thus, within the observed size range of pollination networks ($S = 16$–952 species), testing solely for one kind of network pattern is too simplified (5); both sides of the coin are needed. This lack of correlation between nestedness and modularity suggests that modularity dictates the basic building blocks of networks. These building blocks or their species can then be combined in different contrasting ways, e.g., maximizing either nestedness or modularity. The nested and modular matrix versions are two complementary, yet uncorrelated, analytical steps toward a more profound understanding of network complexity and its causes. In the nested matrix version, species are ranked according to their number of links, whereas in the modular version, they become grouped according to their linkage affinity to other species. This insight, however, could only be gained with the recent appearance of strong algorithms (e.g., 11, 14, 36), allowing a robust identification of modules and a testing of modularity and nestedness. The presence of nestedness in pollination networks tells us that their structure is more than just a string of distinct modules. The link-dense core in the nested version (upper left matrix corner in Fig. 1*A*) is partly composed of between-module links (red cells in Fig. 1*A* and *B*), i.e., links connecting generalist species. The SA

algorithm identifies these generalists (15% of our species sample) by their high z- and/or c-scores. Generalists "glue" peripheral species together into modules, but also modules together into networks and, in this way, blur module boundaries. The key innovative insight obtained from our analysis is that all larger pollination networks are modular and that modularity is complementary to nestedness.

The smallest networks were nonmodular. The reason for this may be a lack of module-detecting power of the algorithm SA when network size decreases. However, the reason may also be related to the connectance of the networks. In general, if connectance in pollination networks increases, then network size decreases (37). If connectance increases, the size of the core of links between generalized species also increases, which may reduce the level of modularity.

Modularity is expected to increase with link specificity (5). It may, for example, be stronger in insect herbivory networks than in pollination and seed-dispersal networks, characterized by low interaction specificity, and in traditional food webs. In an analysis of a herbivory network, flower head-breeding Tephritidae flies and their Asteraceae host species were sorted into six modules (38). Using SA on this network, we detected seven modules and a high modularity (116 species, $M = 0.63$), higher than in most pollination networks ($\bar{M} = 0.52$; see SI Table 1). In this herbivory network, generalists, especially connectors, played a more inferior role as network "glue" (6% of all species) than in pollination networks (15% of all species). Trophic specialization was suggested to cause modularity (38). If so, modularity may be driven by evolution among plant species in various key traits, e.g., chemicals and flower head morphology that "filters out" subsets of species from the overall pattern of species interactions. Reanalyzing a tropical host–parasitoid network also with an expected high link specificity (39), we found an even stronger modularity (104 species, $\bar{M} = 0.67$, 12 modules). Networks of domatia-living ants and their plants are also expected to show high link specificity. Fonseca and Ganade (40) identified seven modules and strong modularity in a South American plant–ant network. Three modules were even isolated from the main network. Using SA, we confirmed their results (41 species, $M = 0.58$). These "long-term intimate mutualisms" (41) strengthen modularity (38, 40, 42), but are rarely seen in networks of interactions among free-living species such as plants and their pollinators or seed dispersers (6).

Many argue for a synthesis of community ecology and evolutionary biology (e.g., refs. 43 and 44). Modularity and nestedness may be the stepping-stone, bridging evolutionary biology and not just community ecology, but also network ecology (45). This may intensify cross-disciplinary work in a research area poor in synthetic field studies (ref. 43; however, see ref. 46). Identification of modules and their component species may be the missing protocol bullets making such studies feasible. If trait evolution within a focal species has only one-to-one species links as its drivers, the network context becomes redundant. However, because such pairwise one-to-one links are extremely rare among free-living mutualists (6, 41) and extremely rare in pollination networks (this study, ref. 47), reciprocal selection is believed to be multispecific, i.e., involving a module, a couple of tightly linked modules, or maybe even an entire small, (non)modular network (6, 44). Within individual modules, connectance was 42%, i.e., almost half of all pollinators and plants were likely to interact directly (Fig. 1B). Thus, the multispecies selection regime in a module must be intense.

Small groups of strongly interacting species are favorite study objects in pollination biology, but without authors explicitly stating that these are modules, e.g., bumblebees and their plants (48) and hummingbirds and heliconias (25, 49). Convergence toward syndromes, e.g., the bird–flower pollination syndrome, in a pollination network may promote modularity (6, 50). Thus, an

identification of modules may bring us closer to an understanding of the basic coevolutionary units that structure complex webs of interaction. This conclusion is strongly supported by the "biological content" of modules. Some of those identified in this study contained a set of species with convergent traits related to their pollination biology, e.g., corolla color or reward type, or they were closely related taxonomically (Fig. 1C). Thus, our study provides a straightforward relationship between a network property and its ecological equivalence expressed as life-history attributes of the component species. The study of modularity may serve to bridge two major coevolutionary research agendas, namely that focusing on small groups of strongly interacting species (6) and that focusing on entire networks (16, 17, 19). In this way, modularity may shed light on how small groups of species scale up to form complex networks of interactions.

The demonstration of modularity in pollination networks has far-reaching conservational implications. Disturbances are expected to spread more slowly through a modular than a nonmodular structure. The identification of the role of species in modular networks may also have valuable bearings to conservation. The network consequences of species extinctions depend on species role, e.g., an extinction of a module hub may cause its module to fragment with no or minor cascading impact on other modules (38), whereas the extinction of connectors may cause the entire network to fragment into isolated modules but with minor impact on the internal structure of individual modules. Conversely, alien invaders of a network, often being highly generalized (51), may cause fusion of modules with profound, long-term effects on network functioning and selection regime.

The omnipresence of modularity and other structural properties, e.g., nestedness, in large pollination networks may change our view on the structuring of biodiversity. Our study shows that modules are small blocks of species, candidating as manageable study objects, and that their study may bridge evolutionary and functional ecology.

Materials and Methods

Database. Pollination networks are a well studied type of networks (e.g., refs. 8, 16–20, and 52–54). We analyzed 51 total pollination networks encompassing almost 10,000 species of plants and flower-visiting animals and 20,000 links, using our own data and data extracted from the literature (see SI Table 1). "Total" does not mean a complete sampling of species and links but only that all plants and flower-visitors observed were included (37). All flower-visiting animals were regarded as pollinators, which, of course, was not necessarily true. Thus, a link is a record of an animal species visiting the flowers of a plant species. The number of links a species has to other species is its degree k. A pollination network is two-modal, i.e., it includes only links connecting plant and animal communities, and it is represented as an adjacency matrix (Fig. 1 A and B; ref. 52), whose elements consist of presence and absence of links (colored and uncolored cells, respectively). Most data are sampled over an extensive period, most often one season, and the sample of study sites covers a wide geographical range and many habitat types (see Table 1). We are aware of the methodological heterogeneity of our database. However, it is more homogeneous than most earlier databases in ecology; for example, almost all taxa were resolved to species level, and links were scored in a similar way, albeit, by many researchers. Data on link strength was not included because such information was lacking in approximately half of the networks and because the module-detecting algorithm (SA) operates only with binary (presence/absence) data (see *SI Appendix*).

Modularity Analysis. A modular network consists of interconnected modules. Each module is a group of species, which are more closely connected to each other than to species in other

modules. To estimate the level of modularity and number of modules in our sample of networks, we used an algorithm (here termed SA) provided by R. Guimerà (for details see refs. 11–13, and see *SI Appendix* for arguments for choice of algorithm). The algorithm, which is based on simulated annealing, identifies modules, whose nodes have the majority of their links inside their own module with an accuracy of 90% (11). Almost all species are unambiguously assigned to a module, except extreme connector species, i.e., species linking equally several modules. For each network, SA calculates an index of modularity M:

$$M = \sum_{s=1}^{N_M} \left(\frac{I_s}{I} - \left(\frac{k_s}{2I} \right)^2 \right), \qquad [1]$$

where N_M is number of modules in the network, I_s is number of links between all species within module s, I is number of links in the network, and k_s is the sum of degrees of all species in s. M values belong to the interval $[0; -1/N_M]$. Basically, M is a measure of the extent to which species have more links within their modules than expected if linkage is random. For each empirical network, we ran an SA-analysis of 100 random networks with the same species degree distribution as the empirical one, and examined whether the empirical network was significantly more modular than the random ones (11). In its calculation of M for a network, the algorithm excludes small isolated modules without any links to the main network. However, we consider these isolates as part of the network. They were identified manually.

Each species was assigned a role by SA according to its topological properties (11). The role of a species is defined by its position compared with other species in its own module and how well it connects to species in other modules. Thus, the role of a species i can be characterized by its standardized within-module degree z:

$$z = \frac{k_{is} - \bar{k}_s}{SD_{ks}}, \qquad [2]$$

and its among-module connectivity, c (identical to ref. 11's participation coefficient P), i.e., how i is positioned with respect to all modules:

$$c = 1 - \sum_{t=1}^{N_M} \left(\frac{k_{it}}{k_i} \right)^2, \qquad [3]$$

where k_{is} is number of links of i to other species in its own module s, $\bar{k}_s$ and SD_{ks} are average and standard deviation of within-module k of all species in s, k_i is degree of species i, and k_{it} is number of links from i to species in module t (including i's own module). If i has all its links within its own module, $c = 0$; and if these are distributed evenly among modules, $c \rightarrow 1$. z and c values are produced by SA, and together they define the roles of species and their position in the zc-parameter space (Fig. 1D). Modifying the criteria of Guimerà and Amaral (11) (who distinguished among seven roles), we sorted all species into peripherals, connectors, module hubs, and network hubs (black, green, blue, and red dots, respectively, in Fig. 2). The latter three are termed generalists. A peripheral species has both a low $z \leq 2.5$ and a low $c \leq 0.62$. It has a few links inside its own module and rarely any to other modules. A connector species has a low $z \leq 2.5$ and a high $c > 0.62$, and it "glues" modules together and is thus important to network coherence. A module hub has a high $z > 2.5$ and a low $c \leq 0.62$ and is important to the coherence of its own module. A network hub has both a high $z > 2.5$ and a high $c > 0.62$ and is thus important to the coherence of both the network and its own module (see Figs. 1D and 2 and refs. 11–13). By using $z = 2.5$ as a cutoff value (11), almost no species with 1–2 links to other species in its own module entered the module hub quadrant ($z > 2.5$, $c \leq 0.62$). The behavior of c was explored by varying the number of links of i to other species in its module (internal links), the number of links of i to other species outside its module (external links), and the number of other modules to which i was linked. With this approach, species with many external links compared with internal links and with links to many modules had $c > 0.62$. In consequence, the cutoff values derived from (11) get a specific topological meaning.

Using the software ANINHADO (36), we estimated level of nestedness for each network, $N = (100 - T)/100$ (16), where T is "matrix temperature." The null model used to assess significance of N was originally provided by ref. 16 and implemented as Ce (36) and the probability that cell a_{ij} has a link is $(a_i/C + a_j/R)/2$ (a_i, number of links in row i; a_j, number of links in column j; C, number of columns; and R, number of rows).

1. Montoya JM, Pimm SL, Solé RV (2006) *Nature* 442:259–264.
2. Pimm SL (1979) *Theor Pop Biol* 16:144–158.
3. Paine RT (1980) *J Anim Ecol* 49:667–685.
4. Pimm SL, Lawton JH (1980) *J Anim Ecol* 49:879–898.
5. Lewinsohn TM, Prado PI, Jordano P, Bascompte J, Olesen JM (2006) *Oikos* 113:174–184.
6. Thompson JN (2005) *The Geographic Mosaic of Coevolution* (Univ of Chicago Press, Chicago).
7. Raffaelli D, Hall SJ (1992) *J Anim Ecol* 61:551–560.
8. Dicks LV, Corbet SA, Pywell RF (2002) *J Anim Ecol* 71:32–43.
9. Girvan M, Newman MEJ (2002) *Proc Natl Acad Sci USA* 99:7821–7826.
10. Roswall M, Bergstrom CT (2007) *Proc Natl Acad Sci USA* 104:7327–7331.
11. Guimerà R, Amaral LAN (2005) *Nature* 433:895–900.
12. Guimerà R, Amaral LAN (2005) *J Stat Mech Theory Exp*, art no P02001.
13. Guimerà R, Mossa S, Turtschi S, Amaral LAN (2005) *Proc Natl Acad Sci USA* 102:7794–7799.
14. Frank KA (1995) *Soc Networks* 17:27–56.
15. Krause AE, Frank KA, Mason DM, Ulanowicz RE, Taylor WW (2003) *Nature* 426:282–285.
16. Bascompte J, Jordano P, Melián CJ, Olesen JM (2003) *Proc Natl Acad Sci USA* 100:9383–9387.
17. Bascompte J, Jordano P, Olesen JM (2006) *Science* 312:431–433.
18. Dupont YL, Hansen DM, Olesen JM (2003) *Ecography* 26:301–310.
19. Vázquez DP, Aizen MA (2004) *Ecology* 85:1251–1257.
20. Jordano P, Bascompte J, Olesen JM (2006) in *Plant–Pollinator Interactions: From Specialization to Generalization*, eds Waser NM, Ollerton J (Univ of Chicago Press, Chicago), pp 173–199.
21. May RM (1972) *Nature* 238:413–414.
22. May RM (1973) *Stability and Complexity in Model Ecosystems* (Princeton Univ Press, Princeton).
23. Faegri K, van der Pijl L (1971) *The Principles of Pollination Ecology* (Pergamon Press, Oxford).
24. Fenster CB, Armbruster WS, Wilson P, Dudash MR (2004) *Annu Rev Ecol Evol Syst* 35:375–403.
25. Stiles FG (1975) *Ecology* 56:285–301.
26. Kato M (2000) *Contr Biol Lab Kyoto Univ* 29:157–252.
27. Bek S (2006) MSc thesis (Univ of Aarhus, Aarhus, Denmark).
28. Arroyo MTK, Primack R, Armesto J (1982) *Am J Bot* 69:82–97.
29. Ingversen TT (2006) MSc thesis (Univ of Aarhus, Aarhus, Denmark).
30. Inoue T, Kato M, Kakutani T, Suka T, Itino T (1990) *Contr Biol Lab Kyoto Univ* 27:377–463.
31. Kato M, Kakutani T, Inoue T, Itino T (1990) *Contr Biol Lab Kyoto Univ* 27:309–375.

ECOLOGY

32. Kato M, Matsumoto M, Kato T (1993) *Contr Biol Lab Kyoto Univ* 28:119–172.
33. Kakutani T, Inoue T, Kato M, Ichihashi H (1990) *Contr Biol Lab Kyoto Univ* 27:465–521.
34. Herrera J (1988) *J Ecol* 76:274–287.
35. Petanidou T (1991) PhD thesis (Aristotelian University, Thessaloniki, Greece).
36. Guimarães PR, Jr, Guimarães P (2006) *Envir Model Softw* 21:1512–1513.
37. Olesen JM, Jordano P (2002) *Ecology* 83:2416–2424.
38. Prado PI, Lewinsohn TM (2004) *J Anim Ecol* 73:1168–1178.
39. Memmott J, Godfray HCJ, Gauld ID (1994) *J Anim Ecol* 63:521–540.
40. Fonseca CR, Ganade G (1996) *J Anim Ecol* 65:339–347.
41. Thompson JN (1994) *The Coevolutionary Process* (Univ of Chicago Press, Chicago).
42. Guimarães PR, Jr, Rico-Gray V, Furtado dos Reis S, Thompson JN (2006) *Proc R Soc London Ser B* 273:2041–2047.
43. Johnson MTJ, Stinchcombe JR (2007) *Tree* 22:250–257.
44. Strauss SY, Irwin RE (2004) *Annu Rev Ecol Evol Syst* 35:435–466.
45. Thompson JN (2006) *Science* 312:372–373.
46. Becerra JX (2007) *Proc Natl Acad Sci USA* 104:7483–7488.
47. Olesen JM, Bascompte J, Elberling H, Jordano P (2008) *Ecology*, in press.
48. Pleasants JM (1980) *Ecology* 61:1446–1459.
49. Temeles EJ, Kress WJ (2003) *Science* 300:630–633.
50. Corbet SA (2000) *Cons Biol* 14:1229–1231.
51. Olesen JM, Eskildsen LI, Venkatasamy S (2002) *Div Distr* 8:181–192.
52. Jordano P, Bascompte J, Olesen JM (2003) *Ecol Lett* 6:69–81.
53. Vázquez DP, Aizen MA (2003) *Ecology* 84:2493–2501.
54. Waser NM, Ollerton J, eds (2006) *Plant–Pollinator Interactions: From Specialization to Generalization* (Univ of Chicago Press, Chicago).

Oleson *et al*. 10.1073/pnas.0706375104.

Supporting Information

Files in this Data Supplement:

SI Appendix
SI Table 1

This Article

▶ **Abstract**

Services

▶ **Email this article to a colleague**
▶ **Alert me to new issues of the journal**
▶ **Request Copyright Permission**

Citing Articles

▶ **Citing Articles via CrossRef**

SI Appendix

Methods to Assess Modularity in Complex Networks. This appendix summarizes salient methods (not exhaustive) and examples to analyze modularity and compartmentalization in complex networks. There is a vast amount of literature on this topic, and the material included here focuses mainly on ecological applications.

Our article uses the modularity analysis method presented by Guimerà and Amaral [Guimerà R, Amaral LAN (2005) *J Stat Mechan*, article no. P02001, pp 1-13.] As described in *Methods* in the main paper, the algorithm maximizes modularity (*M*) using simulated annealing (SA) as a way to simplify the search of combinations. This software is for presence-absence data (now also available for quantitative data in one-mode networks but not available for quantitative data in two-mode networks; R. Guimerà, personal communication). This algorithm is time-consuming, but can handle large datasets. Danon *et al*. [Danon L, Díaz-Guilera A, Duch J, Arenas A (2005) *J Stat Mechan Theory Exp*, P09008] compare different methods to calculate modularity and conclude that the simulated annealing algorithm is the best available. During the review stage of our manuscript, Guimerà *et al*. [Guimerà R, Sales-Pardo M, Amaral LAN (2007) *Phys Rev E* 76: 036102] have provided an extension of this algorithm for two-mode presence-absence networks. Because our plant-animal networks are an example of two-mode networks, one could conclude that this most recent is the algorithm (let's call it SA-2 and refer to the previous one for one-mode networks as SA-1) we should use. However, this decision depends on the sort of question one addresses. R. Guimerà (personal communication) provided the following argument:

First, if one only focuses on one of the interacting communities (animals or plants) the partitions provided by both algorithms are quite similar. Actually, about 75% of the nodes are classified in the same "module" by both algorithms (this computation actually requires to map each module in SA-1 to the best matching module, Guimerà *et al*. (2007). However, SA-1 works with the hypothesis (null model implicit in the definition of modularity) that every node can be connected to every other node, which we know is not true in our networks. It does not invalidate the result but one has to keep this in mind. R. Guimerà (personal communication) does not expect the differences between the two versions of the algorithm to be large.

Second, the two algorithms give slightly different information. SA-1 looks for the most significantly

dense regions (modules) in the network, whereas SA-2 looks for groups of animals (plants) that share the most significantly high number of interactions with the same plants (animals), even when the plants (animals), themselves, belong to different modules. Thus, in a way, the interpretation of SA-2 is harder because groups of plants do not need to correspond to groups of animals. Thus SA-1 is even more interesting in our case because the resulting modules have a precise biological meaning as potential coevolutionary vortices [*sensu* Thompson JN (2005) *The Geographic Mosaic of Coevolution* (Univ of Chicago Press, Chicago)] or as cores of mutualistic networks [*sensu* Bascompte J, Jordano P, Meliám CH, Olesen JM (2003) *Proc Natl Acad Sci USA* 100: 9383-9387]. Also SA-1 is a better choice within a coevolutionary perspective because of the following: SA-2 creates groups of plants somehow independently of the animal groups; i.e., an animal group does not necessarily contain the species that interact with a particular plant group, just animals with a similar pattern of interactions regardless of where the species they interact with are classified in the other set. We are more interested in the coevolutionary units, namely identifying groups of tightly interrelated plants and animals. This is also closer to the concept of pollination syndromes.

The above refers to the modularity analysis based on simulated annealing (SA). There are many other modularity analyses in the literature. Here, we provide a very brief presentation of some of them.

1. **Pimm and Lawton (1980).** This is one of the earliest attempts to characterize food web compartments. It is based on a randomization approach in which a compartmentalization statistic (which appears to be high in compartmented systems) is tested against a null distribution. Species in modules are identified by cluster analysis. It is based on presence-absence data.

   1. Pimm SL, Lawton JH (1980) *J Anim Ecol* 49:879-898.

2. Raffaelli and Hall (1992). The principle of this method is to analyze pairwise trophic similarity by using the Jaccard similarity index. This is not a statistical test, because compartmentalization of a network is not compared with that of randomly assembled networks as in the previous methods. However, it does give an indication of how compartmentalized a network is. It cannot be tested statistically because the similarity measures are non-independent. It is based on presence-absence data. Species in modules are identified by cluster analysis.

1. Raffaelli D, Hall SJ (1992) *J Anim Ecol* 61:551-560.

2. Dicks LV, Corbet SA, Pywell RF (2002) *J Anim Ecol* 71:32-43.

3. Fonseca and Ganade (1996). The log likelihood ratio test statistic is tested by Monte Carlo simulations (random matrices in which row and column totals are fixed). The dataset shows no compartments when using the method of Raffaelli and Hall (1992), because the system has very low connectance (and hence many species pairs have 0 similarity).

1. Fonseca CR, Ganade G (1996) *J Anim Ecol* 65:339-347.

4. Connectivity Correlation [e.g., Melián and Bascompte (2002)]. Connectivity correlation is the correlation between the number of interactions of a species and the average number of interactions of the species it interacts with. This measure was first used in the Internet and protein networks, and later on applied to food webs. A significant negative connectivity correlation characterizes a compartmentalized system. This would correspond to a case where a highly connected node is connected to specialist nodes. Connectivity correlation appears to be a measure of network transitivity (Girvan and Newman 2002, see below) more than modularity, although the two concepts are related. The advantage of this method is that it is based on a simple statistical procedure (a simple correlation) and does not rely on complex software. However, it also provides a simple macroscopic index and does not allow identifying the possible modules.

1. Maslov S, Sneppen K (2002) *Science* 296:910-913.

2. Melian CJ, Bascompte J (2002) *Ecol Lett* 5:705-708.

5. Girvan and Newman (2002). This method is based on an analysis of betweenness. It uses centrality to find compartment boundaries. The algorithm was designed with a sparse network in mind and may not perform well in a dense network (e.g., food webs).

1. Girvan M, Newman MEJ (2002) *Proc Natl Acad Sci USA* 99:   **7821-7826**   .

2. Newman MEJ (2004) *Eur Phys J B* 38:321-330.

3. Newman MEJ (2004) *Phys Rev* E69:066133.

6. Zhou (2003). A new method of hierarchical clustering, using a new dissimilarity measure between vertices to identify "communities" (compartments). Modules can be identified at several different levels of resolution. However, there is no statistical test of which level is most highly significant.

1. Zhou H (2003) *Phys Rev E* 67:061901.

7. Krause *et al.* (2003). This paper is the most convincing evidence of compartments in food webs and illustrates how these network methods can advance our knowledge of ecological questions. Modules are detected using a software first developed by sociologists (Frank 1995), *Kliquefinder*. The number of modules is chosen based on optimization of a criterion, and tested against a null model of no modularity. Odds ratio (interactions within and between modules) is tested by Monte Carlo simulation. This procedure is quite equivalent to the method used in the present paper; both were independently created by two different research fields (sociology and physics).

1. Frank KA (1995) *Social Netw* 17:27-56.

2. Krause AE, Frank KA, Mason DM, Ulanowicz RE, Taylor WW (2003) *Nature* 426:282-285.

8. K-means [Bach and Jordan (2003)]

The K-means is an ordination-based procedure to find groups of components in a multivariate structure. Bach and Jordan discuss a class of algorithms for independent component analysis (ICA) that uses contrast functions based on canonical correlations.

1. Bach FR, Jordan MI (2002) *J Machine Learn Res* 3:1-48.

9. Information-theory based algorithms. This approach uses recent developments in information theory to illustrate and validate an information-theoretic algorithm for module discovery and the resulting measure of network modularity. This measure is an order parameter (a dimensionless number between 0 and 1). Network modularity is quantified by comparison with Monte Carlo generated Erdös-like modular networks. The network information bottleneck (NIB) algorithm is applied to a number of real world networks, such as social networks of coauthorship.

1. Ziv E, Middendorf M, Wiggins CH (2000) *Phys Rev E* 7:046117.

Table 1. Modularity of pollination networks

No.	Habitat, country (ref)	A	P	S	I	N		M		N_M
1	Mixed forest, Japan (1)	840	112	952	1876	0.99	*	0.54	*	10

2	Phrygana, Greece (2)	666	131	797	2933	0.96	*	0.44	*	6
3	Beech forest, Japan (3)	679	91	770	1193	0.99	*	0.63	*	13
4	Scrub/meadow/forest, Amami Isles, Japan (4)	619	107	726	1109	0.98	*	0.64	*	19
5	Ruderal area, Spain (5)	454	31	485	966	0.87	*	0.51	*	8
6	Subalpine forest/meadow, Japan (6)	356	90	446	865	0.97	*	0.54	*	10
7	Urban area, Japan (7)	314	113	427	769	0.98	*	0.55	*	17
8	Grassland, Japan (8)	295	99	394	590	0.98	*	0.64	*	17
9	Forest/grassland, USA (9)	275	96	371	923	0.98	*	0.42	*	12
10	Mosaic heathland, Denmark (YL Dupont unpubl.)	236	33	269	706	0.93	*	0.40	*	5
11	Forest meadow, Denmark (10)	225	37	262	590	0.94	*	0.44	*	5
12	Marsh, Japan (11)	187	63	250	430	0.96	*	0.55	*	10
13	Continuous heathland, Denmark (YL Dupont unpubl.)	186	23	209	468	0.93	*	0.42	*	6
14	Scrub, Spain (12)	179	26	205	412	0.91	*	0.47	*	7
15	Grassland, Cass, New Zealand (13)	153	45	198	403	0.93	*	0.48	*	7
16	Subandean scrub 2200-2600 m, Chile (14)	101	84	185	361	0.96	*	0.51	*	11
17	Subalpine grassland, Craigieburn, New Zealand (13)	130	52	182	383	0.94	*	0.49	*	7
18	Subalpine grassland/scrub, Mount Cook, New Zealand (13)	102	31	133	250	0.94	*	0.49	*	6
19	Tundra, Canada (15)	91	20	111	190	0.95	*	0.46	*	6
20	Cushion zone 2700-3100 m, Chile (14)	64	43	107	196	0.91	*	0.52	*	6
21	Forest 1, Argentina (16, 17)	90	14	104	164	0.91	*	0.48	*	7
22	Coastal scrub, Jamaica (18)	36	61	97	178	0.93	*	0.49	*	7
23	Scrub RB, Argentina	72	23	95	125	0.91	*	0.56	*	10

	(19)									
24	Open forest, Venezuela (20)	46	47	93	151	0.87	*	0.59	*	9
25	Subalpine grassland, Arthur's Pass, New Zealand (13)	68	22	90	147	0.74	*	0.56	*	7
26	Lowland forest, Jamaica (21)	45	29	74	122	0.91	*	0.51	*	7
27	Semi-desert gorge, Tenerife, Spain (22)	51	17	68	130	0.82	*	0.46	*	6
28	Subnival zone 3200-3600 m, Chile (14)	26	41	67	86	0.93	*	0.59	*	7
29	Scrub LD, Argentina (19)	45	21	66	83	0.89	*	0.62	*	6
30	Subarctic rocky slope, Sweden (23)	118	24	142	242	0.86	*	0.51	ns	
31	Scrub/snow gum forest, Australia (24)	81	36	117	253	0.90	*	0.43	ns	
32	Forest, Canada (25)	102	12	114	167	0.93	*	0.47	ns	
33	Tundra, Canada (26)	81	29	110	179	0.95	*	0.48	ns	
34	Wasteground, Denmark (27)	82	26	108	249	0.83	*	0.41	ns	
35	Tundra, Greenland (28)	76	31	107	456	0.74	*	0.25	ns	
36	Meadow, England (29)	79	25	104	299	0.92	*	0.33	ns	
37	Laurel forest, Gomera, Spain (JM Olesen & S Forfang unpubl.)	55	29	84	145	0.95	*	0.46	ns	
38	Lowland forest, Dominica (21)	41	29	70	151	0.88	*	0.40	ns	
39	Forest meadow, Denmark (30)	44	16	60	278	0.78	*	0.21	ns	
40	Tundra, Greenland (31)	39	15	54	92	0.81	*	0.40	ns	
41	Beach-oak forest, Denmark (27)	42	8	50	79	0.59	ns	0.44	ns	
42	Coastal bog, Denmark (27)	40	10	50	72	0.63	ns	0.49	ns	
43	Alpine desert, Tenerife, Spain (32)	38	11	49	108	0.67	*	0.35	ns	
44	Semi-desert slope, Tenerife, Spain (22)	35	14	49	86	0.84	*	0.43	ns	

45	Tundra, Greenland (33)	26	17	43	63	0.85	*	0.46	ns	
46	Deciduous forest, USA (34)	33	7	40	65	0.87	*	0.33	ns	
47	Lowland, Galápagos (35)	22	10	32	27	0.70	*	0.50	ns	
48	Tundra valley, Canada (36)	18	11	29	27	0.78	ns	0.36	ns	
49	Coralline island, Mauritius (37)	13	14	27	52	0.87	*	0.33	ns	
50	Coastal cliff, Azores, Portugal (37)	12	10	22	30	0.67	ns	0.44	ns	
51	Lava desert, Galápagos (38)	6	10	16	16	0.96	ns	0.35	ns	

For networks 1-29, M was significant ($P < 0.05$), and for 30-51 nonsignificant. Each of these two groups was sorted by descending S. A, number of animal species; P, number of plant species; $S = A + P$; I, number of interactions recorded; M, modularity; N, nestedness; *, $P < 0.05$; ns, nonsignificance; N_M, number of modules.

1 Inoue T, Kato M, Kakutani T, Suka T, Itino T (1990) *Contr Biol Lab Kyoto Univ* 27:377-463.

2 Petanidou T (1991) PhD thesis (Aristotelian University, Thessaloniki, Greece).

3 Kato M, Kakutani T, Inoue T, Itino T (1990) *Contr Biol Lab Kyoto Univ* 27:309-375.

4 Kato M (2000) *Contr Biol Lab Kyoto Univ* 29:157-252.

5 González-Álvaro V (2004) PhD thesis (Univ of Castilla-La Mancha, Toledo, Spain).

6 Kato M, Matsumoto M, Kato T (1993) *Contr Biol Lab Kyoto Univ* 28:119-172.

7 Kakutani T, Inoue T, Kato M, Ichihashi H (1990) *Contr Biol Lab Kyoto Univ* 27:465-521.

8 Yamazaki K, Kato M (2003) *Contr Biol Lab Kyoto Univ* 29:255-318.

9 Clements RE, Long FL (1923) *Experimental Pollination. An Outline of the Ecology of Flowers and Insects* (Carnegie Institute of Washington, Washington DC).

10. Bek S (2006) MSc thesis (Univ of Aarhus, Aarhus, Denmark).
11. Kato M, Miura R (1996) *Contr Biol Lab Kyoto Univ* 29:1-48.

12 Herrera J (1988) *J Ecol* 76:274-287.

13 Primack RB (1983) *NZ J Bot* 21:317-333.

14 Arroyo MTK, Primack R, Armesto J (1982) *Am J Bot* 69:82-97.

15 Kevan PG (1970) PhD thesis (University of Alberta, Edmonton, AL, Canada).

16 Vázquez DP, Simberloff D (2002) *Am Nat* 159:606-623.

17 Vázquez DP (2002) PhD thesis (Univ of Tennessee, Knoxville, TN).

18 Percival M (1974) *Biotropica* 6:104-129.

19 Medan D, Montaldo NH, Devoto M, Mantese A, Vasellati V, Bartoloni NH (2002) *Arc Antarc Alp Res* 34:233-241.

20 Ramirez N (1989) *Biotropica* 21:319-330.

21 Ingversen TT (2006) MSc thesis (Univ of Aarhus, Aarhus, Denmark).

22 Stald L (2003) MSc thesis (Univ of Aarhus, Aarhus, Denmark).

23 Elberling H, Olesen JM (1999) *Ecography* 22:314-323.

24 Inouye DW, Pyke GH (1988) *Austr J Ecol* 13:191-210.

25 Helenurm K, Barrett SCH (1987) *Can J Bot* 65:2047-2056.

26 Hocking B (1968) *Oikos* 19:359-388.

27 Montero AC (2005) MSc thesis (Univ of Aarhus, Aarhus, Denmark).

28 Olesen JM, Bascompte J, Elberling H, Jordano P (2008) *Ecology*, in press.

29 Memmott J (1999) *Ecol Lett* 2:276-280.

30 Bundgaard M (2003) MSc thesis (Univ of Aarhus, Aarhus, Denmark).

31 Witt P (1998) BSc thesis (Univ of Aarhus, Aarhus, Denmark).

32 Dupont YL, Hansen DM, Olesen JM (2003) *Ecography* 26:301-310.

33 Lundgren R, Olesen JM (2005) *Arc Antarc Alp Res* 37:514-520.

34 Schemske D, Willson M, Melampy M, Miller L, Verner L, Schemske K, Best L (1978) *Ecology* 59:351-366.

35 McMullen CK (1993) *Pan-Pac Ent* 69:95-106.

36 Mosquin T, Martin JE (1967) *Can Field Nat* 81:201-205.

37 Olesen JM, Eskildsen LI, Venkatasamy S (2002) *Div Distr* 8:181-192.

38 Philipp M, Böcher J, Siegismund HR, Nielsen LR (2006) *Ecography* 29:531-540.