Keystone species and food webs

Ferenc Jordan\textsuperscript{1,2,3,*}

\textsuperscript{1}Collegium Budapest, Institute for Advanced Study, 1014 Budapest, Hungary
\textsuperscript{2}Animal Ecology Research Group, Hungarian Natural History Museum, 1431 Budapest, Hungary
\textsuperscript{3}The Microsoft Research—University of Trento, Centre for Computational and Systems Biology, Piazza Manci 17, 38100 Povo (Trento), Italy

Different species are of different importance in maintaining ecosystem functions in natural communities. Quantitative approaches are needed to identify unusually important or influential, ‘keystone’ species particularly for conservation purposes. Since the importance of some species may largely be the consequence of their rich interaction structure, one possible quantitative approach to identify the most influential species is to study their position in the network of interspecific interactions. In this paper, I discuss the role of network analysis (and centrality indices in particular) in this process and present a new and simple approach to characterizing the interaction structures of each species in a complex network. Understanding the linkage between structure and dynamics is a condition to test the results of topological studies, I briefly overview our current knowledge on this issue. The study of key nodes in networks has become an increasingly general interest in several disciplines: I will discuss some parallels. Finally, I will argue that conservation biology needs to devote more attention to identify and conserve keystone species and relatively less attention to rarity.

Keywords: centrality; food web; indirect effect; keystone species; network analysis

1. INTRODUCTION

We live in the age of the sixth mass extinction. Extinction rates may reach one to three species per hour (May \textit{et al.} 1995); a rate probably never seen before. We know that biodiversity is important, but we do not yet clearly understand its functional aspects and the possible ways to maintain it (Terborgh 1999, p. 100). To understand the functional diversity, we would need to know more about the roles particular species play in ecological communities (Jones \& Lawton 1995; Kareiva \& Levin 2003). The identification of key players (and dummies) on the Hutchinson stage must be quantitative in order to be the objective and predictive. This is the short-term task of the research for keystone species (Power \textit{et al.} 1996); the final aim might ultimately be to outline a conservation policy based primarily on functional importance, rather than on rarity. Conserving biodiversity is often a compromise between protecting species, areas, ecosystems or processes (Simberloff 1998). If our goal is to conserve ecosystem structure and maintain a reliable supply of ecosystem services, then identifying and conserving important species will be the only real long-term solution for managing diversity loss, ultimately protecting only the rarest species is only a symptomatic treatment.

There are several kinds of ‘focal’ species in ecology and conservation biology: A variety of ecological studies have identified endemic, dominant, link, indicator, invasive, introduced, reintroduced and keystone species (see also Dale \& Beyeler 2001); conservation biologists also consider umbrella, charismatic and flagship species (Simberloff 1998). A question to ask is ‘how do these concepts overlap and how operative is their use?’ (Dale \& Beyeler 2001). All of these concepts are probably useful for some purposes, but it is the very diversity of these terms that makes objective studies more difficult. All of them acknowledge that there is a need for clarifying the most important species in ecosystems. Although, we should also delineate a difference between species that are important for nature (e.g. for maintaining ecosystem functions) and species maybe more important for scientists (e.g. Red List species of limited ecological roles): several species of high aesthetic or economic value are certainly important to be protected, but it might be admitted that they are protected for other reasons than maintaining ecosystem functioning.

It is not easy to define ‘importance’ from an ecological perspective. A fairly widespread view is to expect a species to be more important, if its extinction (or disturbance) will have a larger effect on the abundance of others in the community. This influence may be sufficiently large so as to lead to secondary extinctions. For example, ecosystem engineers have a large impact on other species as they drastically modify the abiotic environment (Jones \textit{et al.} 1994); if an ‘ecosystem engineer’ is disturbed, many other species will also be influenced. However, species more typically influence their partners via biotic, interspecific interactions (e.g. predation, facilitation) and any impact on one species will typically have an immediate effect on its partners. For example, a decreasing number of pollinators may have a negative effect on the plant species that they interact with (Kearns \textit{et al.} 1998). From this point of view, we can define important species as major interactors that have many links to others. Their extinction may break coevolved relationships among coadjusted sets of species; and this may significantly harm ecosystem integrity.
In this paper, (i) I discuss the major biological problems related to the position of species in interaction networks, (ii) I review and propose some quantitative tools of network analysis that are helpful in characterizing the position and positional importance of species in ecological networks, and (iii) I overview our recent knowledge about the relationship between network position and community dynamics. Finally, I discuss the relevance of the problem to future conservation practices.

2. KEY POSITIONS IN FOOD WEBS

Studying food webs (and ecological interaction networks in general) provides a synthetic framework for representing, analysing and visualizing the rich network of interspecific interactions in ecosystems (Cohen 1978; Pimm 1982, 1991; Dunne 2006). The possible effects that spread out from loss or dramatic reduction in abundance of a species may easily be traced via its predator–prey, or other kinds of direct, pairwise, interactions including pollination (Jordano 1987), competition (Paine 1984) or ‘facilitation’ (e.g. providing shelter, Bruno et al. 2003).

All of these interactions are likely to work simultaneously (e.g. facilitation and competition, Bertness & Shumway 1993) and their combinations give rise to well-defined indirect interactions such as the trophic cascades that result from the loss of a keystone predator (Menge 1995; Abrams et al. 1996). These, mostly two, and three step long indirect effects enable biotic effects to cascade through the network of interactions. Indirect effects are sometimes considered weak and meaningless, but they can also be long and strong. The effect of Alaskan fisheries on Californian sea otters is an example that has both spatial, as well as topologically long effects (i.e. number of links in the food web, Estes et al. 1998).

Similarly, the overall effect of Iberian lynx on rabbits provides an example where the indirect effect (trophic cascade) is stronger between two species than the direct one (predation): lynx positively influences rabbit abundance by feeding on mesopredators that potentially have a much larger impact on rabbit abundance (Palomares et al. 1995). Thus, indirect interactions may be quite long and relatively strong.

Direct partners and the indirect neighbourhood define the position of a species (or a trophic group) in the ecological-interaction network. There are food webs where the positional importance of certain species (or functional groups) is of central importance. The best examples are probably the ‘wasp-waist’ ecosystems of pelagic upwelling zones (Cury et al. 2000). Here, a large number of species at low and high trophic levels are linked by a single, or only a few, species in the middle trophic levels. These may be sardines, anchovies, jellyfish, krill or copepods (Bakun 1996). For example, the importance of krill in subarctic waters seems to be a clear consequence of its special position in the network (Ducklow et al. 2007; Smith et al. 2007). Wasp-waist species are often not only major interactors, but also the key regulators of both higher and lower trophic levels: their behaviour may switch between alternative energy pathways (Chavez et al. 2003; Stibor et al. 2004; Murphy et al. 2007). They also behave as major energy gates (Margalef 1968); and their strange dynamics may be a simple consequence of their particular network position (Jordán et al. 2005).

3. QUANTIFYING POSITIONAL IMPORTANCE

(a) The network perspective

If indirect links between species may be long and strong, we may want to consider them when assessing how intimately a particular species is linked within the interaction network. Network analysis provides a variety of tools for characterizing the topological importance of nodes in graphs (representing species in communities). Certain methods have been imported into ecology from sociology (Harary 1961; Jordán et al. 1999; Krause et al. 2003) and economics (Hannon 1973) this has led to an emerging methodological synthesis (Jordán & Scheuring 2004; Pascual & Dunne 2006; Estrada 2007). Local, mesoscale and global structures are characterized by various network indices for neighbourhoods of various sizes. In this paper, we only discuss local and mesoscale indices, as we are primarily interested in assessing species’ key roles in the system, rather than in characterizing whole systems by global network measures.

As soon as species can be quantitatively ranked by their topological importance, then keystone species may be quantitatively identified, and the keystone term will no longer be restricted to a ‘yes or no’ quality. This is a key condition of predictability. Topological key-stone species may be used as quantitative indicators and may be helpful in setting objective conservation priorities. But the rich system of interspecific interactions calls for approaches that better understand their internal structure. Statistical physics provides macroscopic descriptions of large and complex networks, but it reduces ecological information to a few numbers (for example, the diameter of a network, see Williams et al. 2002).

There are basically three options for better understanding the biology of these networks. First, components of a large network may be aggregated so that they constitute a smaller web that is easier to analyse. This is a standard technique in ecosystems ecology (e.g. Christensen & Pauly 1992), but the bias in biological content is potentially large and hard to understand (there are no clear aggregation rules and significant taxonomic biases). Second, an isolated subgraph of the interaction network may be studied, as occurs in the case of plant–pollinator (Jordano 1987; Memmott 1999; Bascompte et al. 2003; Waser & Ollerton 2006) and host–parasitoid (Memmott et al. 1994; Müller et al. 1999; Rott & Godfray 2000; Lewis et al. 2002) networks. In this case, there are always significant problems concerning how the neglected external effects are on members of the subgraph (e.g. predators of pollinators). Nevertheless, details of the network remain accurate and may provide important insights about robustness and redundancy. A third option is to consider the heterogeneity of the network, to determine critical nodes and to focus research efforts on them (figure 1). This is the method outlined and explored in this paper. However, it must be noted that all of these approaches overlap: for example every ecological network is aggregated to some extent.
The number of neighbours—a local view
The most straightforward way to initially proceed to potentially identify keystone species in networks is to use the degree, $D$, number of neighbours of a graph node to assess its positional importance. Degree and its distribution (link distribution) have already been used extensively in food-web studies (e.g. Dunne et al. 2002; Montoya & Solé 2002). This is the most local measure and often provides a fast and simple evaluation. However, based on our knowledge of indirect effects, the results may be misleading: key species may be important interactors even if they only have a small number of direct partners in the web. For example, Pisaster ochraceus, the prototype of keystone species has a rich indirect interaction system, but only a few important direct partners (Paine 1969). Thus, a wider neighbourhood of graph nodes needs to be considered both conceptually and practically (Jordán & Scheuring 2002; Brose et al. 2005).

Neighbours of neighbours—a mesoscale view
Mesoscale indices quantifying the network position of nodes in graphs are frequently based on distance, i.e. the number of links on the shortest path connecting nodes $i$ and $j$. The two most widely used indices have been imported from social network analysis (Wasserman & Faust 1994); one is closeness centrality (measuring the average distance of the focal node from all others in the graph), while the other is betweenness centrality (the proportion of the shortest paths between all $i$ and $j$ pair of nodes that contains the focal node). These have recently been introduced to ecology (Estrada 2007; Jordán et al. 2007) as well as many other disciplines.

The interest in quantifying indirect relationships between nodes in ecological networks is not new. The first application must have taken place by Harary (1961, ‘net status’ of species), modified later by Jordán et al. (1999, keystone index) in ecology. These applications consider directed graphs, and the application of ‘net status’ was the first case when species were ranked according to their quantified importance (as early as in 1961). A related approach is to quantify dominator trees in binary (Allesina & Bodini 2004), and weighted (Allesina et al. 2006), directed food webs, providing information on the structurally most important nodes.

In biological control studies, the strength of apparent competition between host species in host–parasitoid communities required quantitative assessment. Godfray and colleagues (Godfray et al. 1999; Müller & Godfray 1999; Müller et al. 1999; Rott & Godfray 2000, Lewis et al. 2002; Van Veen et al. 2006) have introduced and elaborated a simple technique for quantifying and visualizing this (the quantitative parasitoid overlap graph). Following the same logic, their method has been adapted to a more general use; namely to also quantifying longer indirect interactions (Jordán et al. 2003a). This index enables one to analyse indirect interactions of various lengths separately (up to a three-step length). It assumes a network with undirected links where interspecific effects may spread in any direction without bias (we are interested in interaction webs, in the...
Here, chemically and behaviourally mediated indirect effects are not considered, only indirect chain effects (Wootton 1994). The effect of species $j$ on species $i$, when $i$ may be reached from $j$ in $n$ steps, is defined as $a_{n,ij}$. The simplest case of calculating $a_{n,ij}$ is when $n = 1$ (i.e. the effect of $j$ on $i$ is direct): $a_{1,ij} = 1/D_i$, where $D_i$ is the degree of node $i$ (i.e. the number of its direct neighbours including both prey or predatory species). We assume that indirect chain effects are multiplicative and additive (figure 2). When the effect of step $n$ is considered, we define the effect received by species $i$ from all $N$ species in the same network as:

$$\psi_{n,i} = \sum_{j=1}^{N} a_{n,ij},$$

which is equal to 1 (i.e. each species is affected by the same unit effect.). Furthermore, we define the $n$-step effect originated from a species $i$ as

$$\sigma_{n,i} = \sum_{j=1}^{N} a_{n,ij},$$

which may vary among different species (i.e. effects originated from different species may be different). Here, we define the topological importance of species $i$, when effects up to $n$ steps are considered as

$$\text{TT}_n = \frac{\sum_{m=1}^{n} \sigma_{m,i}}{n} = \frac{\sum_{m=1}^{n} \sum_{j=1}^{N} a_{m,ij}}{n},$$

which is simply the sum of effects originated from species $i$ up to $n$ steps (one plus two plus three...up to $n$) averaged over by the maximum number of steps considered ($n$). We note that this index may be calculated also for weighted networks (e.g. Scotti et al. 2007); similarly to the original one where parasitation frequency data were widely available (Godfray et al. 1999).

By this TT$^*$ index, it is possible to quantify the origins of effects influencing a particular species, i.e. the internal interaction structure of the network. For simplicity, I illustrate this on the relatively small food web of the Kuosheng Bay, Taiwan (Lin et al. 2004). In figure 3a, the size of nodes is proportional to the relative strength of effects originated from node 2 (black node). It is obvious that nodes 7 and 11 are mostly influenced, while the right side of the network is only weakly affected. In figure 3b–d, the same is illustrated for nodes 9, 12 and 15, respectively. Note that node 12 has a larger effect on node 11 (its indirect partner) than on node 4 (a direct partner), because of the particular topology of the network (figure 3c). Also, trophic cascades coming from the top-predator (node 15) are stronger for producer 2 than on the others (nodes 8 and 1, figure 3d). The summarized TT$^*$ values for each species are represented by the sizes of nodes in figure 3e: based on this index, node 14 is the most important species (functional group) in this community (see numerical values in table 1). Note that the importance rank of species depends on $n$ (the maximum length of indirect effects): for $n=10$, species 4 is the most important interactor in the ecosystem (table 1). It is also clear that degree ($D$) predicts very similar, but somewhat different results (note also the ties, table 1); differences may be much larger in case of larger networks. For more details on various centrality indices, see Jordán & Scheuring (2004) and Estrada (2007) and for more on the relationships between different indices, see Friedkin (1991) and Jordán et al. (2007).

(d) Overlapping neighbourhoods

The vector of $a_{n,ij}$-values for species $j$ has been defined as its ‘trophic field’ (Jordán 2001). For long indirect effects, every species is connected to every other. It is reasonable to define a $t$ threshold of $a_{n,ij}$-values separating strong interactive partners from weak interactors. Given a maximum length of indirect effects ($n$) and a threshold for interaction strength ($t$), every node may be characterized by its effective trophic range (Jordán et al. in press). Since the sets of strong interactors of two, or more, nodes may overlap, it is important to quantify the positional uniqueness of

**Figure 2. Calculating the TI-index.** (a) the strength of a direct effect is the reciprocal of the degree ($D$) of the affected node ($a_{1,CD}$ stands for the topological strength of the one-step effect of node $D$ on node $C$). (b) For indirect effects, we suppose multiplicative (b) and additive (c) effects.
graph nodes. The ‘trophic field overlap’ (TO\textsubscript{i,j}) between nodes i and j is the number of strong interactors appearing in both i’s and j’s effective range. The sum of all TO-values between species i and others (∑TO\textsubscript{i,j} summed over all j with i≠j) provides the summed trophic field overlap of species i (TO\textsubscript{i}). This may be normalized by dividing it with the maximum value (TO\textsubscript{max}) for a given network (reTO\textsubscript{r} = TO\textsubscript{i} / TO\textsubscript{max}). Note that all this is determined by i, n and the topology of the network. We define the ‘topological uniqueness’ of species i as TU\textsubscript{i} = 1 – reTO\textsubscript{r} and its ‘trophic function’ as TF\textsubscript{i} = TI\textsubscript{i} + TU\textsubscript{i}. This index indicates keystone species by considering both the importance (TI) and the uniqueness (TU) of a position. We have calculated the TF\textsubscript{5.0.1} values for each species (table 1, far right) and illustrated the results in figure 3f: the size of nodes is proportional to TF\textsubscript{5.0.1}. Note that node 9 ranks 2, while its best rank has been 7 so far. So, species 9 has a position of average importance, but outstanding uniqueness.
This approach reflects an early definition of keystone species (single-species functional group, see Bond 1994). In fact, a really important species is expected to be hardly replaceable in the community. In a network context, this means a relatively unique interaction structure within the network. For how TF depends on $n$ and $t$, see Jordán et al. (in press).

Functional redundancy and the overlapping role of species have already been considered in previous studies in somewhat different ways (Luczkovich et al. 2003; Shannon & Cury 2003). We understand that this approach is far more interesting for larger webs; this small network was only used for illustration.

4. FOOD WEB POSITION AND COMMUNITY EFFECT

Studying network topology is more interesting and useful, if we see the link between structures and dynamics (see Pimm (1980) for one of the first approaches). Local and mesoscale indices of nodes correlate with their simulated dynamics in some small model food webs. For example, both the degree and the orientation of links (whether linking to preys or to predators) do influence extinction risk in a simple model of coupled Lotka-Volterra (LV) equations describing population dynamics (Jordán et al. 2002). Based on similar LV models, Eklof & Ebenman (2006) have found that degree and, especially, the number of consumers of deleted species influences extinction within the community. Beyond topological centrality, a lower trophic position of deleted nodes also appears to be influential for triggering extinctions (Borrvell et al. 2000; Ebenman et al. 2004; Eklof & Ebenman 2006), as well as the skewed distribution of interaction strength values (Borrvell et al. 2000) and particular arrangements of strong links (Jonsson et al. 2006). Most importantly, the indirect interaction pattern is also influential (Brose et al. 2005; Ebenman & Jonsson 2005).

In larger food webs, modelled by more complicated dynamics, it has been demonstrated that mesoscale indices may better predict extinction than does the local degree (Quince et al. 2005). Brose et al. (2005) have shown that there is a marked range of indirect effects in complex networks influencing local dynamics. Based on a somewhat different simulation approach (EwE mass-balance models, Christensen & Pauly 1992), indirect neighbourhood has been shown to best correlate with simulated behaviour (Libralato et al. 2006). Comparing simulated mass-balance effects of disturbing species (Okey 2004) to their local and mesoscale topological properties, in both binary, and weighted, trophic networks, suggests that it is weighted indices (also considering indirect effects) that best correlate with simulated importance (Jordán et al. 2008). It is important to emphasize this, because many food web studies focus on node degree in binary networks (i.e. a local, unweighted description, Solé & Montoya 2001), where the correlation pattern between various structural indices and dynamics characterizes the relevance of degree-based, local approaches.

5. CONCLUSIONS

In most communities, several key species play a relatively large role as they have many links to others in the interspecific interaction network. Their network position may be characterized by mesoscale centrality indices beyond the local measure of degree, i.e. the number of neighbours. These metrics consider the neighborhood of each species and provide a more realistic evaluation of how rich the interaction pattern of a particular species is. This approach is particularly important in systems where indirect effects dominate. Beyond quantifying the positional importance of single graph nodes, it is also of interest to what extent these trophic fields overlap. The overlap measure quantifies topological redundancy in interaction networks and identifies nodes that can be characterized with a rich, as well as unique interaction pattern. The significance of having a unique pattern of interactions depends, among other things, on the switching ability of species. If switching is generally easy and the food web has a flexible structure, it is less interesting to identify unique positions. On the contrary, if species are less opportunistic consumers, organisms with unique interactions may really be candidate key species. In case of highly aggregated networks, links among large trophic groups are less flexible, so measuring the overlap may be more important.

Recent simulations generally support the view that the indirect interaction pattern cannot be neglected when looking for topological keystone species. The importance of indirect effects, well documented in the field, is quantifiable by network analysis. Better comprehension of the relevant range of indirect effects may help food web research to become a more practical and more applicable science (Winemiller & Layman 2005). It is noted that most of the results discussed here are equally relevant for both food webs (in a strict sense) and interaction networks (in a broader sense).

The mesoscale view outlined in this paper emphasizes that, although theoretically everything is connected to everything else in a network, it is possible in practice to map and quantify the relative strengths of connections and dependencies. Important tasks for future studies are: (i) to incorporate the switching effect into dynamical models (Pelletier 2000; Jordán et al. 2003b), (ii) to analyse the phenology of ecological interaction networks (descriptive time-series of food webs), (iii) to improve our knowledge on the exact biological meanings of different centrality indices, (iv) to better fit adequate methods to actual problems, and (v) to improve the comparative analysis of methodologically similar food webs (Baird et al. 1991). Still, this field may provide the greatest help in rendering conservation biology more quantitative and predictive, as well as function- and system-based.

There is a current major interest in key nodes in many kinds of networks ranging from animal social networks (Wey et al. 2008) through landscape graphs (Pascual-Hortal & Saura 2006) to transport networks (Jordán 2008a). Studying central nodes in networks appears to be a way to better understand complex systems in a relatively easy way. The comparative view on several kinds of systems (Faust & Skvoretz 2002) may promote mutual understanding.

*Phil. Trans. R. Soc. B* (2009)
and application even to societal problems (McMahon et al. 2001; Jordán 2008b). Cross-fertilizing methods (betweenness centrality pervades every discipline), especially if well adapted, make scientific progress very fast and efficient.

Current simulations propose that studying important network positions may be a key to better understand community dynamics, so structurally important species may need particular attention in future conservation practice. For example, copepods or earthworms may be functionally much more important than rare, sexy vertebrates and their protection may have tremendous direct and indirect positive effects on ecosystem integrity and functioning (and indirectly help rare species, too). How to set conservation priorities is an evergreen question (Mace & Collar 2002): future conservation biology should probably focus more on ‘the little things that run the world’ (Wilson 1987) — if they are in special network position.

I am grateful to Thomas Okey, Wei-chung Liu, Tim Wyatt, Simone Libralato, Marco Scotti, János Podani, István Schéming, Zsófia Benedek, Vera Vasas, Barbara Bauer and Agnes Mike for their help and cooperation. I am grateful to Kati Csepi for checking the English and the anonymous reviewers for their useful comments. My research was fully supported by Society in Science: The Branco Weiss Fellowship, ETH Zürich, Switzerland. I also thank the organizers of the Port Elizabeth and San Jose meetings for their invitation.

REFERENCES


Bakun, A. 1996 Patterns in the ocean: ocean processes and marine population dynamics. San Diego, CA: University of California Sea Grant Program.


Jorda´n, F. 2001 Trophic fields.


Pascual-Hortal, L. & Saura, S. 2006 Comparison and development of new graph-based landscape connectivity...


