A robust measure of food web intervality

Daniel B. Stouffer*, Juan Camacho**, and Luis A. Nunes Amaral***

*Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL 60208; and **Departament de Física (Física Estadística), Universitat Autònoma de Barcelona, Bellaterra, 08193 Catalonia, Spain

Edited by Robert May, University of Oxford, Oxford, United Kingdom, and approved October 11, 2006 (received for review May 10, 2006)

Intervality of a food web is related to the number of trophic dimensions characterizing the niches in a community. We introduce here a mathematically robust measure for food web intervality. It has previously been noted that empirical food webs are not strictly interval; however, upon comparison to suitable null hypotheses, we conclude that empirical food webs actually do exhibit a strong bias toward contiguity of prey, that is, toward intervality. Further, our results strongly suggest that empirically observed species and their diets can be mapped onto a single dimension. This finding validates a critical assumption in the recently proposed static niche model and provides guidance for ongoing efforts to develop dynamic models of ecosystems.

D

epite their complexity, the structure of natural food webs displays a number of remarkable regularities (1–9). The existence of these empirical regularities has prompted several researchers to develop simple models that aim to identify the mechanisms that underlie food webs. In particular, three recent “static” models, the niche model (2), the nested-hierarchy model (7), and the generalized cascade model (9) predict key statistical properties of food webs from a variety of environments, including deserts, rain forests, lakes, and estuaries.

Stouffer et al. (9) demonstrated that these three models share two fundamental mechanisms that account for the models’ success in reproducing many of the empirical patterns. (i) Species form a totally ordered set in niche space, that is, species can be ordered along a single niche dimension. (ii) Each species has an exponentially decaying probability of preying on a given fraction of the species with equal or lower niche values (9). Despite these similarities, the models exhibit some differences; a crucial distinction concerns how species’ prey are organized along the single dimension. In the niche model, species prey on a contiguous range of prey. In the nested-hierarchy and generalized cascade models, in contrast, the diets are not restricted to a contiguous range.

The differences in prey selection lead to drastic differences in the intervality of the food web graph6 (Fig. 1 a and b). The significance of intervality in complex food webs was first noted by Cohen (10), who reported, as did subsequent studies (1, 11–13), that the vast majority of empirical food webs in the literature appeared to be interval graphs. Significantly, these studies also suggested that the probability that a food web is interval strongly depends on the number of species represented in the food web, decreasing from approximately one for very small food webs to close to zero for larger webs (1). The food webs that were analyzed in these studies typically comprised very few species, leaving open the question of whether, or to what degree, larger and more complex food webs are interval (1). More recent studies reported persistent nonintervality of highly resolved empirical food webs (2, 7).

Importantly, the degree of intervality of a food web is related to the number of trophic dimensions characterizing the possible niches in a community (11). More specifically, one may ask what is the minimum number of variables required to describe the factors that influence the trophic organization of the species in a community? Is this number the same or different for different communities (14–16)? If a food web is interval, then the species and their diets can be represented along a single dimension. It has been suggested that a single factor (species’ mass) provides a suitable proxy for this dimension (6, 17–21). Any departure from intervality has been understood to imply additional complexity in the mechanisms responsible for the structure of the food web.

The number of higher-quality food web data sets has steadily increased, and these data have enabled researchers to uncover a number of solid empirical regularities (2–5, 7, 9). Thus, we believe that a more definitive answer to the question of food web intervality may be at hand.

In this article we address the question of how “noninterval” empirical food webs truly are. To this end, we define a measure of intervality that is more robust than those already in the literature. Notably, our results agree with previous studies that observed that empirical food webs are strictly noninterval; however, we demonstrate that their degree of “intervality” can be understood as a perturbation on an underlying interval structure. Our results provide support to the conjecture that species and their diets, that is, ecosystem niches, may be mapped onto a single dimension.

Food Web Intervality

In the studies of Cohen et al. (1), Cohen (10, 11), and Sugihara (12, 13), intervality was reported as a binary variable: a web either “is” or “is not” interval. Recently, two local estimates have been used to measure the “level of diet discontinuity” (7). The first measure, $D$$_{int}$, is defined as the number of triplets of species with an “irreducible gap” divided by the number of possible triplets. An irreducible gap is a gap in a consumer’s diet that cannot be made contiguous because of the constraints imposed by other consumers’ diets (Fig. 1).

The second measure, $C$$_{int}$, is defined as the number of chordless cycles of length four in the consumer overlap graph. In the consumer overlap graph, two consumers are connected if they share at least one prey. That is, if species $A$ and $B$ share prey with species $C$ and $D$, the consumer overlap graph would consist of links $A$ $\leftrightarrow$ $C$, $A$ $\leftrightarrow$ $D$, $B$ $\leftrightarrow$ $C$, and $B$ $\leftrightarrow$ $D$. This is a cycle because it is possible to travel from any one of the four species to any other in this graph. If species $A$ and $B$ do not share any prey and similarly species $C$ and $D$ do not share any prey either, this cycle is “chordless” and the four diets cannot be made contiguous simultaneously (11). The measure $C$$_{int}$ is related to Sugihara’s
If we consider a noninterval food web and attempt to reproduce here in detail. In the idealized case of a fully interval food web, diets are organized along a single dimension.

This process yields an ordering of species in the food web in such a way as to generate the deviation from an interval food web. This is one of potentially multiple permutations that can give rise to the same value of $G$. In this particular case, one can find an ordering with $G = 0$. It should be noted that

(12) rigid circuit property, which states that in an interval food web every circuitous path of length $l \geq 4$ in the consumer overlap graph is shortened by a chord.

Using these two measures, Cattin et al. (7) reported that the nonintervality of empirical food webs is a significant food web pattern. Caution, however, is required because both $D_{dct}$ and $Cy_4$ yield only local estimates of intervality and cannot be directly extrapolated to an entire ecosystem.

Specifically, a cycle of length four in the consumer overlap graph with a chord may still contain irreducible gaps (1). Therefore, $Cy_4$ is, at best, a lower bound for what Cattin et al. (7) intended to measure. Likewise, when computing $D_{dct}$, the normalization factor used by Cattin et al. (7) accounts only for multiphagous consumers, not all species. By concentrating on species triplets, the resulting measure is an overestimation and not amenable to comparisons between food webs of different sizes and linkage densities. Moreover, as pointed out by Martinez et al. (22), Cattin et al. (7) also do not address what values of $D_{dct}$ or $Cy_4$ would in fact be statistically significant or represent a large deviation from an interval food web.

In contrast to previous studies, we determine here the degree of intervality of an entire food web. To do this, we first find the order of species in the food web in such a way as to generate the “most interval” ordering of the food web. This process yields the best approximation to a food web where the species and their diets are organized along a single dimension.

We discuss our definition of most interval and its justification here in detail. In the idealized case of a fully interval food web, each consumer’s diet is represented by a single contiguous range. If we consider a noninterval food web and attempt to reproduce the idealized web as closely as possible, we will want all prey of a given predator to “appear” as close together as possible on the resource axis (Fig. 1). For example, for a given consumer, a sequence of two adjacent prey, a gap of one species, and two more adjacent prey (i.e., $\ldots PP PP \ldots$), where $P$ represents a prey and $-$ represents a nonprey) is preferable to the same sequence but with a gap of two species or larger (e.g., $\ldots PP PP \ldots$). Indeed, the former situation would be far more likely given an interval web that experienced random omissions or changes, such as those possibly introduced by field sampling.

For a food web graph $G$ with $S$ species, there are $S!$ possible species orderings $O_k(S) = s_1, s_2, \ldots, s_S$ with $k = 1, \ldots, S!$. Because of the large number of possible permutations, it is computationally infeasible to determine the best ordering through enumeration. It is for this reason that we use simulated annealing, a heuristic technique that significantly reduces the computational effort required to find an optimal or close-to-optimal solution (see Methods and ref. 23 for details).

When attempting to find the most interval ordering, the objective is to minimize the discontinuity of all predators’ prey (Fig. 1). We thus define a cost function $G(O_k)$, which is the sum of the gaps in all consumers’ diets:

$$G(O_k) = \sum_{i=1}^{S} \sum_{j=1}^{n_i} (g_i^j)^p.$$  \hspace{1cm} [1]

Here $n_i$ is the number of gaps in the diet of species $i$ and $g_i^j$ is the number of species in the $j$th gap in the diet of species $i$ for a given
ordering $O_k(T)$. Here we report results for $\beta = 1$; however, the selection of other values, such as $\beta = 2$, yields similar orderings of the empirical data (Supporting Text and Fig. 3, which are published as supporting information on the PNAS web site). Simulated annealing yields an estimate $\hat{G}$ for the total number of gaps $G = \min \{ G_k(\mathcal{G}) \}$ of the food web. The smaller $\hat{G}$ is the more interval the food web is.

**Null Hypotheses for Food Web Intervality**

As happens for other graph and combinatorial problems, the actual value of $G$ is of little significance (24); rather, one needs to assess whether the measured value of $G$ is significantly different from the expected value under suitable null hypotheses. To solve this problem, we have designed three complementary null hypotheses that place different restrictions on how consumers' diets may be organized within a food web.

Our first null hypothesis is the set of randomizations of the empirical food web. We perform this randomization by using the Markov-chain Monte Carlo switching algorithm (25, 26) and treat single, double, and cannibal links separately (see Methods for details). The randomized empirical food web stands as a food web graph with no constraints placed on consumers' diets. That is, in the randomization there is no correlation between the prey of a given species and their organization on the resource axis. We therefore expect that $\hat{G}$ for these randomized food webs will be maximal. Comparison to this null hypothesis thus provides verification of whether there are any structural regularities in the organization of species' diets within empirical food webs.

Our second null hypothesis is the set of food webs generated by the generalized cascade model (9). In the generalized cascade model, each of the $S$ species $i$ are assigned a niche value $n_i$ drawn from a uniform distribution in the interval $[0,1]$. A predator $j$ selects at random a fraction $x$ of the species $i$ with niche values $n_i \leq n_j$ as its prey, where $x$ is drawn from a $\beta$-distribution $p(x) = \beta(1 - x)^{(\beta - 1)}$. Here $\beta = (S^2/2L) - 1$ and $L$ is the number of trophic links in the ecosystem.

The generalized cascade model food webs are generated with the same number of species $S$ and linkage density $L/S$ as the empirical food webs. Whereas randomization of the empirical food webs imposes no structural constraints on consumers' diets, the generalized cascade model does. Each predator may then select their prey from random, but instead of from the entire resource axis, their selections are restricted to only those species with niche values lower than or equal to their own. This mechanism leads to a smaller number of gaps for species placed lower on the resource axis. Comparison of the empirical data to this null hypothesis will provide evidence as to whether empirically observed diets exhibit additional structural constraints.

Comparison to the two previous null hypotheses will provide an indication of whether empirical food webs have a larger number of gaps than would be expected for random structures with little or no bias toward contiguity of prey. To quantify any bias toward contiguity of prey in empirical food webs, we need to develop a third null hypothesis, which we base on a generalization of the niche model (2).

Let us first recall the definition of the niche model. Each of the $S$ species $i$ are assigned a niche value $n_i$ drawn from a uniform distribution in the interval $[0,1]$. A predator $j$ in the niche model preys on a range $\tau_j = n_j$ of the resource axis, where $x$ is drawn from a $\beta$-distribution just as in the generalized cascade model. The center of the range $\tau_j$ is selected uniformly at random in the interval $[\tau_j/2, \tau_j]$. All species $i$ whose niche values $n_i$ fall within this range are considered prey of species $j$.

To allow for a tunable bias toward prey contiguity, we generalize the niche model in the following manner. First, we reduce the range $\tau_j$ for a predator $j$ to $\tau'_j = cr_j = cn_j x$, where $c$ is a fixed parameter in the interval $[0,1]$. Because species are distributed uniformly at random on the resource axis, a predator $j$ with range $\tau_j$ has on average $n_S \tau$ prey. The same applies to the reduced range $\tau'_j$, and therefore a predator has $\Delta k = (\tau_j - \tau'_j) S = (1 - c) r_j S$ expected prey unaccounted for after the range reduction. Next, we select these $\Delta k$ prey, rounded to the nearest integer value, randomly from the species $i$ with niche value $n_i \leq \tau'_j$.

Table 1. Comparison of empirical data with the random model and the generalized cascade model

<table>
<thead>
<tr>
<th>Food web</th>
<th>$\mathcal{G}_{\text{e}}$</th>
<th>$\langle \mathcal{G}_R \rangle$</th>
<th>$z_R$</th>
<th>$p_R$</th>
<th>$\langle \mathcal{G}_{GC} \rangle$</th>
<th>$z_{GC}$</th>
<th>$p_{GC}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benguela</td>
<td>27</td>
<td>81</td>
<td>-10.31</td>
<td>$&lt; 10^{-10}$</td>
<td>78</td>
<td>-3.59</td>
<td>$&lt; 10^{-3}$</td>
</tr>
<tr>
<td>Bridge Brook Lake</td>
<td>1</td>
<td>51</td>
<td>-11.24</td>
<td>$&lt; 10^{-10}$</td>
<td>48</td>
<td>-4.42</td>
<td>$&lt; 10^{-5}$</td>
</tr>
<tr>
<td>Canton Creek</td>
<td>615</td>
<td>810</td>
<td>-7.65</td>
<td>$&lt; 10^{-10}$</td>
<td>1,804</td>
<td>-9.38</td>
<td>$&lt; 10^{-10}$</td>
</tr>
<tr>
<td>Caribbean Reef</td>
<td>298</td>
<td>498</td>
<td>-12.47</td>
<td>$&lt; 10^{-10}$</td>
<td>340</td>
<td>-1.17</td>
<td>0.12</td>
</tr>
<tr>
<td>Chesapeake Bay</td>
<td>11</td>
<td>48</td>
<td>-5.86</td>
<td>$&lt; 10^{-8}$</td>
<td>38</td>
<td>-3.08</td>
<td>$&lt; 10^{-2}$</td>
</tr>
<tr>
<td>Coachella Valley</td>
<td>51</td>
<td>117</td>
<td>-10.74</td>
<td>$&lt; 10^{-10}$</td>
<td>64</td>
<td>-1.04</td>
<td>0.15</td>
</tr>
<tr>
<td>Grassland</td>
<td>5</td>
<td>28</td>
<td>-5.08</td>
<td>$&lt; 10^{-6}$</td>
<td>95</td>
<td>-5.76</td>
<td>$&lt; 10^{-8}$</td>
</tr>
<tr>
<td>Little Rock Lake</td>
<td>427</td>
<td>1,347</td>
<td>-26.75</td>
<td>$&lt; 10^{-10}$</td>
<td>1,641</td>
<td>-9.89</td>
<td>$&lt; 10^{-10}$</td>
</tr>
<tr>
<td>Northeast U.S. Shelf</td>
<td>700</td>
<td>1,291</td>
<td>-17.76</td>
<td>$&lt; 10^{-10}$</td>
<td>1,050</td>
<td>-5.84</td>
<td>$&lt; 10^{-8}$</td>
</tr>
<tr>
<td>St. Marks</td>
<td>157</td>
<td>343</td>
<td>-14.12</td>
<td>$&lt; 10^{-10}$</td>
<td>258</td>
<td>-2.95</td>
<td>$&lt; 10^{-2}$</td>
</tr>
<tr>
<td>St. Martin</td>
<td>95</td>
<td>204</td>
<td>-12.18</td>
<td>$&lt; 10^{-10}$</td>
<td>193</td>
<td>-4.06</td>
<td>$&lt; 10^{-4}$</td>
</tr>
<tr>
<td>Scotch Broom</td>
<td>23</td>
<td>226</td>
<td>-15.67</td>
<td>$&lt; 10^{-10}$</td>
<td>508</td>
<td>-8.94</td>
<td>$&lt; 10^{-10}$</td>
</tr>
<tr>
<td>Skipwith Pond</td>
<td>26</td>
<td>36</td>
<td>-3.32</td>
<td>$&lt; 10^{-3}$</td>
<td>42</td>
<td>-1.59</td>
<td>0.06</td>
</tr>
<tr>
<td>Stony Stream</td>
<td>645</td>
<td>915</td>
<td>-9.82</td>
<td>$&lt; 10^{-10}$</td>
<td>2,225</td>
<td>-12.56</td>
<td>$&lt; 10^{-10}$</td>
</tr>
<tr>
<td>Ythan</td>
<td>270</td>
<td>513</td>
<td>-11.69</td>
<td>$&lt; 10^{-10}$</td>
<td>915</td>
<td>-8.52</td>
<td>$&lt; 10^{-10}$</td>
</tr>
</tbody>
</table>

*Note that we use $G$ to refer to the actual minimum number of gaps for the most interval ordering of a food web, whereas $\mathcal{G}$ refers to the estimate obtained with simulated annealing. The only case when we can be certain that $G = \mathcal{G}$ is when $G = 0$. 

Stouffer et al.
Empirical Results

We study 15 empirical food webs from a variety of environments: three estuarine, Chesapeake Bay (27), St. Marks (28), and Ythan (34); five freshwater, Bridge Brook Lake (30), Canton Creek (31), Little Rock Lake (32), Skipwith Pond (33), and Stony Stream (31); three marine, Benguela (34), Caribbean Reef (35), and Northeast U.S. Shelf (36); and four terrestrial, Coachella Valley (37), Grassland (38), Scotch Broom (39), and St. Martin (40) (Table 3, which contains supporting information on the PNAS web site).

For each empirical food web, we obtain \( G_c \) (Table 1). We find \( 1 \leq G_c \leq 700 \) for all food web, that is, none of the webs is interval. To compare these empirical values to our three null hypotheses, we perform the following steps. For each empirical food web, we generate a minimum of 100 model food webs corresponding to the respective null hypothesis and obtain \( G_{\text{model}} \) for each model food web.

We then want to be able to estimate the probability that the value \( G_c \) appears given each null hypothesis. To do this, we examine not just the mean of \( G_{\text{model}} \) but its probability distribution (Fig. 2). Using the Kolmogorov-Smirnov test (41), we determine that we cannot reject the hypothesis that the \( G_{\text{model}} \) values are drawn from a Gaussian distribution. We then use the fact that a Gaussian distribution describes the model data to directly calculate an estimate for the probability of observing a value of \( G_{\text{model}} \approx G_c \).

We first compare the set of empirical food webs \( \{ T \} \) to the set of randomized food webs \( \{ T_R \} \) (Table 2). We find that, for every food web, \( G_c < G_R \). To estimate the significance of this difference for each of the individual food webs, we calculate the probability that the model exhibits a value \( G_{\text{GC}} \leq G_c \). For 12 of the 15 food webs, \( p_R < 10^{-10} \). For the remaining three food webs, \( p_R < 10^{-3} \).

We now compare the set of empirical food webs to the set of generalized cascade model-generated food webs \( \{ T_{GC} \} \) (Table 1). We again find that for every empirical food web, \( G_c < G_{GC} \). We find that for 12 of the 15 food webs the probability that \( G_{GC} \leq G_c \) is again quite small, \( p_{GC} < 10^{-2} \). For the remaining three food webs, Skipwith Pond, Coachella Valley, and Caribbean Reef, we find larger probability values, 0.06, 0.12, and 0.15, respectively. Further analysis indicates that \( p_{GC} \) decreases with the directed connectance \( L/S^2 \) for a fixed number of species \( S \). Thus the higher values of \( p_{GC} \) are likely caused by the higher connectance of these webs.

To conclusively reject the two random hypotheses, we apply a Bonferroni correction (42), which decreases the significance

\[ n_j \] that are not already a prey of species \( j \). If \( c = 0 \), we recover the generalized cascade model, whereas for \( c = 1 \), we recover the niche model.

For densely connected food webs, predators typically have greater numbers of prey. Because these prey are constrained to have a niche value less than or equal to the predators, the greater the directed connectance the greater the probability that these prey are contiguous, despite the random predation. This is more pronounced for smaller than for larger food webs.
Discussion

The concept of “niche theory” or “niche space” is a fundamental concept in the study of ecosystems. Niche space was classically defined as an “n-dimensional hyperspace” with n given by the innumerable ecological and environmental characteristics (14, 15). Therefore, each species’ niche is the “result” of all n factors acting on it and the niche represents the functional role and position of the organism in its community. The more recent “interpretation” of niche theory, however, relates to the niche providing species an ordering or hierarchy (15, 18, 20). This formulation provides a much simpler criterion than Hutchinson’s (14) “n-dimensional hyperspace.” Studies have suggested that by using species’ mass or size a food web can in fact be mapped to a single dimension (6, 18–21, 43). Furthermore, the placing of species onto a single dimension is a crucial ingredient in many models developed to describe food web structure (1, 2, 7, 9).

Recently, however, discussions as to how interval food webs truly are, were renewed by the contrast between the niche model, and its contiguous range of prey, and the generalized cascade and nested-hierarchy models, and their random predation (9). Our results allow us to conclusively demonstrate that natural ecosystems, while not fully interval, are significantly more interval than would be expected by chance alone. Indeed, we find the empirical food webs to be statistically indistinguishable from model food webs whose diets are between 85% and 100% contiguous. The idea that species and their diets can be so closely mapped onto a single dimension represents a significant insight that can guide us on how best to go about developing dynamic ecosystem models such as the recent integration of the niche model and nonlinear bioenergetic modeling proposed by Martinez et al. (44).

A number of future questions must be answered before the topic of food web intervality can come to a close. First and foremost is getting a better understanding of exactly what processes are behind the deviations from truly interval behavior. While some of the gaps within species diets may be caused by interactions not observed during field sampling, we find it unlikely that all gaps may be attributed to this factor. It was noted earlier, albeit on different food webs from those studied here, that ecosystems with multiple habitats, for example, an estuary, are less likely to be interval than single-habitat food webs (1, 10). Indeed, one would not expect food webs containing several habitats to be strictly interval because each habitat is likely to have its own separate resource axis.

It would likewise be very interesting to examine additional properties of the “most-interval” ordering or orderings, \( \{O_k\} \) (Supporting Data Set, which is published as supporting information on the PNAS web site). Studies that compared these orderings to those obtained when comparing species’ masses, or related properties (6), would be particularly intriguing.

Methods

Simulated Annealing. Simulated annealing is a stochastic optimization technique that enables one to find a “low-cost” configuration while still broadly exploring the space of possibilities (23). This is achieved by introducing a computational “temperature” \( T \). When \( T \) is high, the system can explore configurations of high cost, whereas at low \( T \) the system can only explore low-cost regions. By starting at high \( T \) and slowly decreasing \( T \), the system descends gradually toward deep minima.

For each iteration in the simulated annealing algorithm, we attempt to swap the position of two randomly selected species to go from the initial ordering \( O(T) \) to the proposed ordering \( O(T) \). This updated ordering \( O(T) \) is then accepted with probability

\[
\frac{e^{-E/T}}{1 + e^{-E/T}}
\]

where \( E = E - E^* \), and \( E^* \) is the lower bound of \( E \) for a particular ordering.

**It should be noted that our results may exhibit some underestimation of c in particular as noted earlier for densely connected food webs such as Coachella Valley, Northeast U.S. Shelf, and Skipwith Pond.
\[
P = \begin{cases} 
1 & \text{if } G(O_i) \leq G(O_f) \\
\exp \left( -\frac{G(O_f) - G(O_i)}{T} \right) & \text{if } G(O_i) > G(O_f) 
\end{cases} \quad [4]
\]

where \(G(O_i)\) is the cost after the update and \(G(O_f)\) is the cost before the update. For each value of \(T\), we attempt \(qS^2\) random swaps with \(q \geq 250\). After the movements are evaluated at a certain \(T\), the system is “cooled down” to \(T = cT\), with \(c = 0.99\).

**Generating Randomized Networks.** To generate an ensemble of random networks, one must first define the constraints of the randomization (45, 46). In our analysis, we preserve the following attributes for each species during randomization of the food web: (i) number of prey, (ii) number of predators, (iii) number of single links, \(A \rightarrow B\), (iv) number of double links, \(A \leftrightarrow B\), and (v) whether or not a species is a cannibal.

We use the Markov-chain Monte Carlo switching algorithm (26) and treat single, double, and cannibal links separately. For example, two single links \(A \rightarrow B\) and \(C \rightarrow D\) become \(A \rightarrow D\) and \(C \rightarrow B\), provided both \(A \rightarrow D\) and \(C \rightarrow B\) do not already exist in the network and they do not form new double links. Similarly, two double links \(A \leftrightarrow B\) and \(C \leftrightarrow D\) become \(A \leftrightarrow D\) and \(C \leftrightarrow B\), provided that both \(A, D\) and \(C, B\) are unconnected by a link in any direction.

We thank R. Guimerà, S. Levin, R. D. Malmgren, C. A. Ng, M. Sales-Pardo, E. N. Sawardecker, and M. J. Stringer for stimulating discussions and helpful suggestions. D.B.S was supported by National Science Foundation-Integrative Graduate Education and Research Traineeship “Dynamics of Complex Systems in Science and Engineering” Grant DGE-9987577. J.C. was supported by Comisión Intermínisterial de Ciencia y Tecnología Grants BFM2003-06033 and SGR00186 and the Departament d’Universitats, Recerca i Societat de la Informació of the Generalitat de Catalunya. L.A.N.A was supported by a National Institute of General Medical Sciences/National Institutes of Health K25 Career Award, the J. S. McDonnell Foundation, and the W. M. Keck Foundation.