Syst. Biol. 61(2):240–252, 2012

DOI:10.1093/sysbio/syr098
Advance Access publication on September 14, 2011

Towards a Dynamic Analysis of Weighted Networks in Biogeography

DANIEL A. DOS SANTOS[1,2], MARÍA GABRIELA CUEZZO[1,2], MARÍA CELINA REYNAGA[2], AND EDUARDO DOMÍNGUEZ[1,2,*]

[1]*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET); and* [2]*Instituto de Biodiversidad Neotropical, Facultad de Ciencias Naturales e Instituto Miguel Lillo, Universidad Nacional de Tucumán, Miguel Lillo 205 (CP 4000), Tucumán, Argentina;*
[*]*Correspondence to be sent to: Instituto de Biodiversidad Neotropical, Facultad de Ciencias Naturales e Instituto Miguel Lillo, Universidad Nacional de Tucumán, Miguel Lillo 205 (CP 4000), Tucumán, Argentina; E-mail: mayfly@unt.edu.ar.*

Received 25 February 2011; reviews returned 11 May 2011; accepted 5 July 2011
Associate Editor: Adrian Paterson

Abstract.—An improvement to the Network Analysis Method (NAM) in Biogeography based on weighted inference and dynamic exploration of sympatry networks is proposed. Intricate distributions of species result in a reticulated structure of spatial associations. Species are geographically connected through sympatry links forming an overall natural network in biogeography. Spatial records are the signals that provide evidence to infer these sympatry links in the network. Punctual data are independent of a priori area determination. NAM is oriented to detect groups of species embedded into the global network that are internally sustained by sympatric cohesiveness but weakly connected (or disconnected) to outgroup entities. These groups, called units of co-occurrence (UCs), are segregated through the iterative removal of intermediary species according to their betweenness scores. Instances of analysis of the original NAM are improved through the following changes and extensions: (i) inference of weighted sympatry networks using new measures sensitive to the strength of overlap and topological resemblance between set of points; (ii) construction of a basal network discriminating major from minor sympatry associations; (iii) evaluation of the entire process of iterative removal of intermediary species for the selection of UCs found on different subnetworks; (iv) network partitioning based on the intrinsic cohesiveness of the UCs; (v) production of a graphical tool (cleavogram) depicting the structural changes of the network along the removal process. Improvements are tested using real and hypothetical data sets. Resolution of patterns is notably increased due to a more accurate recognition of allopatric patterns and the possibility of segregating spatially overlapped UCs. As in original NAM, spatial expressions of UCs are building blocks for biogeography supported by strictly endemic and connected species through sympatry paths. [Clusters pattern recognition; cohesiveness; dot maps; NAM; spatial point process; sympatry.]

Networks are a collection of elements (nodes or vertices) connected by some relationships of interest (links or edges). Networked structures arise in a wide array of different contexts such as technological and transportation infrastructures, social phenomena, and biological systems (Barrat et al. 2004). Thinking of them as networks and studying their patterns of connection can often lead to new and useful insights (Newman 2010). Species connected through their sympatry links constitute a natural network in biogeography. Two major issues are associated to this approach: the inference of sympatry links and the identification of clusters within the overall network. Dos Santos et al. (2008) developed the Network Analysis Method (NAM) to address those issues. In considering the inference of sympatric links, NAM focuses in the use of direct evidence provided by species records (punctual data) and differentiates from the traditional procedure where sympatry is derived by overlapping a priori delimited species areas. Once the sympatry network is obtained, NAM identifies clusters of cohesively sympatric species that are simultaneously allopatric with others. These entities, called units of co-occurrence (UC), are obtained through the removal of intermediary species between those groups. Betweenness score (Freeman 1977) indicates the intermediacy level of each taxon because it measures the occurrence frequency of a taxon in a geodesic path connecting two other taxa. Therefore, the removal of species with high betweenness disaggregates the network. Segregating clusters display different levels of cohesiveness. Group cohesiveness is the force that keeps group members together, so as a group approaches a clique configuration with every individual species tied directly to every other individual species, its cohesiveness consolidates. Clustering coefficient (Watts and Strogatz 1998) denotes the tendency of a group to form a clique. NAM was based on the performance of the betweenness and clustering coefficient to identify UCs. The spatial or geographical expressions of each UC consisted of all the records known for each species and represent candidates to areas of endemism. In a spatial context, species are strictly endemic to each UC. The final status of these candidates will depend on the congruent historical relationships of the involved taxa (Humphries and Parenti 1999; Mast and Nyffeler 2003; Dos Santos et al. 2008; Parenti and Ebach 2009).

NAM as originally presented needed some improvements in the following subjects: (i) a way to measure the strength of sympatric associations between species to discriminate different degrees of overlap; (ii) the achievement of a more accurate sympatry inference avoiding assumptions that overestimate species ranges based on punctual data; (iii) the generation of a network partitioning based on the cohesiveness of species groups; (iv) a graphical tool to show the successive subnetworks obtained during the iterative removal of intermediary species. The first two items refer to the sympatry matrix generation, whereas the last two correspond to the analysis itself. Casagranda et al. (2009) criticized the operational procedure of interpenetration criterion and the use of binary relationships between taxa not considering different degrees of overlap, showing that these can lead to resolution problems in

the generation of the sympatry matrix. Items 1 and 2 address these criticisms.

The goal of this paper was to present improvements that increase NAM resolution and applicability. In order to address Items 1–2, a weighted sympatry inference is developed using the minimum spanning tree (MST) as our basic tool to evaluate spatial occupancy. Then, the network analysis is performed to recover the UCs according to their inherent cohesiveness (Item 3). Finally, a new graphical tool, called a *cleavogram*, is provided to summarize the total process of iterative removal of intermediary species (Item 4). The performance of these improvements is tested using both empirical and hypothetical examples.

Methodological Improvements and Theoretical Background

Fundamentals and Definitions

Informally, networks are objects composed of elements and interactions or connections between those elements. The natural means to model networks mathematically is supplied by the notion of graphs. A graph $G = (V, E)$ is an abstract object formed by a set V of vertices (nodes) and a set E of edges (links) that connect pairs of vertices. Nodes represent elements of the network and edges correspond to their interactions. The two vertices joined by an edge are called its endvertices. If two vertices are connected through a link, they are adjacent and we call them neighbors. Numerical values (weights) can be associated to the different edges of the graph and are useful to describe the strength of association. In dealing with sympatry networks, weights help to differentiate strong (or coextensive) from weak (or minor) species range overlap. Edge weights can be represented as a function $w : E \to \mathbb{R}$ that assigns each edge $e \in E$ a weight $w(e)$. An unweighted graph is equivalent to a weighted graph with unit edge weights for all $e \in E$, that is, $w(e) = 1$. A graph $G' = (V', E')$ is a subgraph of the graph $G = (V, E)$ if $V' \in V$ and $E' \in E$. It is a vertex-induced subgraph if E' contains all edges $e \in$ E that joins vertices in V′. If C is a proper subset of V, then $G - C$ denotes the graph obtained from G by deleting all vertices in C and their incident edges. In this text, a given sympatry subnetwork obtained through removal of intermediary species is equivalent to a vertex-induced subgraph. The operation itself corresponds to the previous subtraction $G - C$ where the target network and the intermediary species (with their incident edges) are G and C, respectively. An undirected graph is connected if every vertex can be reached from every other vertex, that is, if there is a path from every vertex to every other vertex. A graph consisting of a single vertex is also taken to be connected. A path from $s \in V$ to $t \in V$ is an alternating sequence of nodes and edges, beginning with s and ending with t, such that each edge connects its preceding node with its succeeding node. The path length is the sum of the weights of its edges. The minimum length of any path connecting a pair of nodes is called geodesic distance. Graphs that are not connected are called disconnected. For a given undirected graph $G = (V, E)$, a connected component of G is an induced subgraph $G' = (V', E')$ that is connected and maximal (i.e., there is no connected subgraph $G'' = (V'', E'')$ with $V'' \supset V'$). A clique in a graph consists of a subset of nodes, all of which are adjacent to each other where there are no other nodes that are also adjacent to all the members of the clique. Ultimately, the different UCs that NAM pursues to identify are connected components of the various subnetworks emerging along the removal process. For more details on network analysis and algorithms, see Wasserman and Faust (1999) and Brandes and Erlebach (2005).

Inference Instance

Ideally, two facts should be reflected in measuring the strength of spatial association between species ranges: the amount of overlap and their overall topological resemblance. For example, suppose we have three species A, B, and C distributed in the following points A = {1, 2}, B = {1, 3}, and C = {1, 4}, the amount of overlap is the same for any couple of species, that is, point 1, so they are equally similar according to this sole element. However, if we know that points 2 and 3 are closer than any of them with point 4, we can state that species A and B are more related than either of them are to C. Usually, spatial similarity has been estimated as a function of the shared elements between lists of operational geographic units (OGUs) occupied by species, despite the fact that this procedure entails an arbitrary grid demarcation and discounts topological relationships among OGUs. To avoid these drawbacks, we will focus on punctual records and consider their overall pattern of spatial occupancy.

Let S_i and S_j be point sets associated to species i and j with cardinalities $|S_i|$ and $|S_j|$, respectively. If both species i and j share all their records $S_i \cup S_j = S_i = S_j$, then both species are geographically linked by coextensive sympatry. However, identically distributed species are hardly found on real punctual data, and a measurement of the deviation of data from the case of maximal coextensive sympatry (homopatry) is necessary. The general procedure to estimate that deviation consists of studying the area occupancy induced by each species and identify then the amount of change obtained after combining point sets S_i and S_j into a single undifferentiated set $S_{ij} = S_i \cup S_j$.

Two different strategies based on this general procedure are used to obtain a weighted matrix of spatial association between species.

Strategy based on geometrical layout.—The spatial coverage induced by each species can be estimated from the length of the MST associated to its point set. A MST on a set of points is defined as the shortest network interconnecting the given points with all edges between the points. For a set of points in the plane, locations are given in Cartesian coordinates, and the respective MST

consists of straight lines connecting the points so that the sum of their length is minimized (Euclidean MST). In a geographic context, locations of the points are defined by their latitude and longitude. The standard metric in the sphere, which corresponds to the L_2 metric in Euclidean space, is the orthodromic or great circle distance. The MST projected over the curved surface of earth consists then of arcs of great circle between the points so that the overall length is minimized (Geodesic MST). Prim's (1957) algorithm can be used to construct the MST. This algorithm is valid whether the distances between the points come from the Euclidean distance or the orthodromic distance on the sphere (Dolan et al. 1991).

Given the spatial coverage associated to species i and j: MST (S_i) and MST (S_j), respectively, the pooled set S_{ij} is expected not to increase that coverage if S_i andS_j are spatially codistributed. So, sympatry is suggested when MST $(S_{ij}) \leq$ MST (S_i) + $MST(S_j)$ because this inequality reflects no gain in the space occupancy after the union of sets. The following expressions retrieve values analogous to the three classical parameters a, b, and c used in similarity indices calculations (Legendre and Legendre 1998) but meaning here either shared (a) or unique (b and c) spatial information associated to the pair of species under comparison:

$a = \text{MST}\ (S_i) + \text{MST}(S_j) - \text{MST}\ (S_{ij}),$

$$b = \begin{cases} 0, & \text{if MST}(S_j) \leq a \\ \text{MST}\ (S_i), & \text{if } a \leq 0 \\ \text{MST}\ (S_i) - a, & \text{otherwise,} \end{cases}$$

$$c = \begin{cases} 0, & \text{if MST}\ (S_j) \leq a \\ \text{MST}\ (S_j), & \text{if } a \leq 0 \\ \text{MST}\ (S_j) - a, & \text{otherwise.} \end{cases}$$

Finally, the strength of sympatric association is calculated through the positive matching index (Dos Santos and Deutsch 2010) based on the average fraction of shared information along the continuous interval ranging from the lowest to the largest item under comparison:

$$\frac{1}{(a + \max(b,c)) - (a + \min(b,c))} \int_{a+\min(b,c)}^{a+\max(b,c)}$$

$$\frac{a}{x}\mathrm{d}x = \frac{a}{|b-c|} \ln\left(\frac{a + \max(b,c)}{a + \min(b,c)}\right).$$

When items are equally sized, $a + b = a + c$, this index must be calculated as $a/(a + b) = a/(a + c)$. Figure 1 helps to visualize the behavior of this measure as two different dot clouds gradually approximate and interpenetrate. This formula is used for cases of presumptive overlap between patterns of spatial coverage (i.e., $a \geqslant 0$, $a + b + c > 0$) and constrains the sympatric association to the interval [0, 1]. The particular case of two rare species

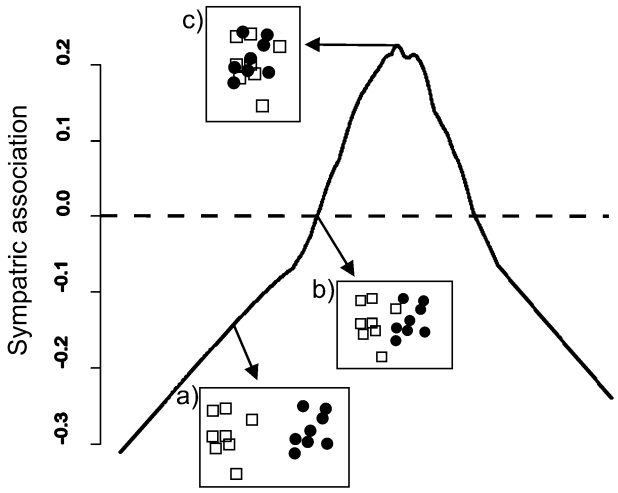

FIGURE 1. Coefficient of sympatric association as function of dot clouds proximity and interpenetration. Three scenarios are illustrated: a) allopatry, b) minimal sympatry strength, and c) maximal sympatry strength.

co-occurring in their single record is considered as evidence for strict sympatry, and the respective entry in the adjacency matrix is set to one. Finally, if the comparison between two sets of points returns some $a < 0$, then |a| should be interpreted as the shortest interspecific gap between those sets measured in the units of a.

Strategy based on nearest interspecific records.—A second procedure to study spatial association between point sets consists of the measurement of the cost to convert one set of points into the other and vice versa so that complete overlap is achieved. This procedure can be used in conjunction with or as an alternative to the previous strategy. This cost is evaluated through the shortest distance separating each point from the species not occurring in it (Fig. 2). The shorter the distance for covering unoccupied points the higher the congruence between distributions. If both species co-occurred in a point, the cost would be zero for that point. The average cost across the joint set of points can be used to summarize the degree of spatial association into a single score. The average cost for spatial homogenization (ACSH) for two species i and j is formalized as follows:

$$\text{ACSH} = \frac{1}{n}\sum_{k=1}^{n} \max\{d_{kS_i}, d_{kS_j}\},$$

where n denotes the size of the pooled set of records (i.e., $|S_i \cup S_j|$), k indexes each of the points under study, d_{kS_i} and d_{kS_j} correspond to the smallest geographical distance between the point k and the set of points of species i and j, respectively. When punctual records are represented by cell grids, the distance would correspond to the Euclidean separation between them using their row and column numbers as coordinates.

The above equation weights the points equally a reasonable approach if points are regularly spread out over a given region. However, as in real data, the points

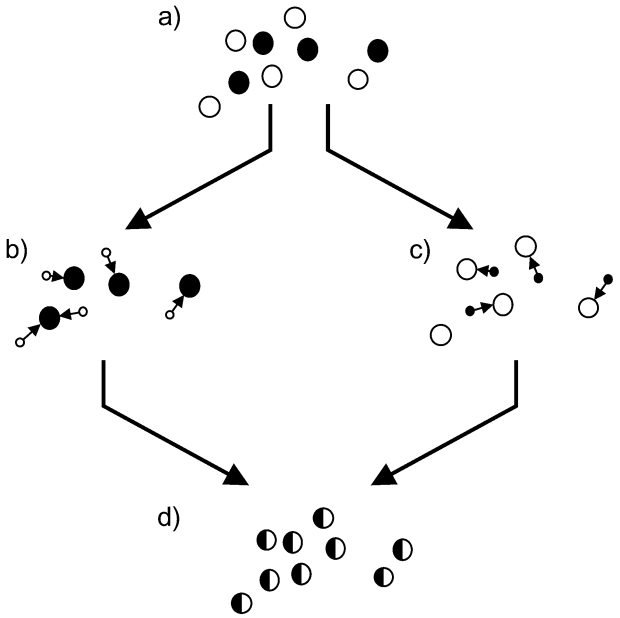

FIGURE 2. Logic underlying ACSH calculation. (a) Original records of a pair of species. (b and c) Nearest interspecific distance for each record. (d) Ideal scenario of complete overlap between species.

can be frequently clumped due to either natural patchiness or unequal sampling efforts; the use of differential weights is a more realistic approach. Thus, many records lying in close proximity but associated to a tiny fraction of the entire species range should receive less weight than a few records widely scattered throughout the same species range. Otherwise, the redundant spatial information associated to the former scenario could locally bias the ACSH result. For each species point set, the weight assigned to each point is proportional to the average length of MST arcs incident on it. Weights are then normalized so that they sum to unity. Additionally, it is useful to study the codistribution in groups with more than two species. As very proximate and interpenetrated species are successively analyzed for grouping, it is expected that the measure will remain with a low scoring and relatively stable. Consequently, the general weighted ACSH for two or more taxa is defined as:

$$\mathrm{ACSH}' = \frac{1}{S}\sum_{k=1}^{n} w_k \max\{d_{kS_i}, d_{kS_j}, \ldots, d_{kS_q}\},$$

where w_k reflects the normalized weight associated to point k and S counts the number of involved species. For any point common to more than one species, its weight is updated by adding its respective values individually estimated in each species point set.

Analysis of the Matrix

The inference instance yields weighted matrices of spatial association between species that can be treated separately or together. Weighted matrices are dichotomized using a threshold to generate a binary matrix corresponding to the basal network to be analyzed by NAM. Here, the process of dichotomization is described for the case of topological matrices, obtained under the first strategy.

Let $\mathbf{A} = [a_{ij}]$ the adjacency matrix of spatial association based on geometrical layout, where each element indicates allopatry ($a_{ij} < 0$) or different strength of sympatric association ($0 \leq a_{ij} \leqslant 1$) between species i and j.This matrix is dichotomized through threshold t obtaining the binary matrix $\mathbf{B} = [b_{ij}]$ such as $b_{ij} = 1$ if $a_{ij} \geqslant t$, otherwise $b_{ij} = 0$. For sympatry networks, t could fall in the range from 0 to 1. As a general rule, the level of interpenetration and proximity between species records increases as t approaches 1. However, the choice of a single threshold is unconcerned with the relative importance of the edge weights. To preserve the information content associated with the weighted network, edges with high relative importance are retained. The relative importance of a link is determined by its relative weight with respect to all links incident to involved endvertices.

The scores of matrix **A** are truncated via two thresholds (t_1, $t_2 | t_1 \leq t_2$), so that entries of **A** lower than t_1 become -1, whereas entries equal to or higher than t_2 are set to 1. By default, $t_1 = 0$ and $t_2 = 1$. A credit for relative strength is added to each edge when its weight (i) is not negative and (ii) is not exceeded by the weights of other links incident to its endvertices. Given that $a_{ij} \geq 0$, a reweighted matrix $\mathbf{R} = [r_{ij}]$ is then derived from **A** through $r_{ij} = S + 1/(2 + L)$, where S denotes the number of times $a_{ij} \geq a_{ik}$ AND $a_{ij} \geq a_{jk}$ and L refers to the frequency of observing $a_{ij} < a_{ik}$ AND $a_{ij} < a_{jk}$, for all $k \neq i, j$ belonging to the set V. The entry r_{ij} of **R** reflects the number of triads in which the edge between species i and j has the highest score. It is also inversely rewarded by the number of triads in which that edge has the lowest score.

Note that matrix **A** is transformed into another matrix with only two entries, that is, -1 and 1, when $t_1 = t_2$. Edges below t_1 (recoded -1) will be 0 at the final binary matrix because edges cannot receive credits of strength. On the other hand, edges at or above t_2 (recoded 1) will be one at the final binary matrix because they have the theoretical maximum score. In this way, this represents the simplest dichotomization of weighted matrix through a single threshold.

Network partitioning and cleavogram.—NAM is oriented to identify groups of species that meet the requirement of within-group sympatry and between-group allopatry. These groups of species correspond to UCs in a subnetwork. The different UCs are usually embedded into a more global network due to intermediary species connecting allopatric groups. Then, the removal of intermediary species will segregate the different UCs. In Dos Santos et al. (2008), only one of the resulting subnetworks was selected. A positive increase in the overall clustering performance favored one subnetwork over other suboptimal ones. A problem with this criterion was that the information associated with the entire removal process was not available for further analysis.

A new strategy for network partitioning based on individual group cohesiveness rather than clustering properties of the whole subnetwork is implemented. As a group approaches a clique configuration, its cohesiveness tends to the maximum. However, in real sympatry networks, missing links are expected due to several causes such as differential spatial occupancy, poor sampling, taxonomic misidentifications, etc. Thus, the cohesiveness concept must consider structural holes in the networks (missing links). The following parameters are used to assess cohesiveness in our sympatry networks: (i) density: proportion of all the theoretically possible edges between nodes that actually exist; (ii) graph diameter: length $\max_{u,v} d(u,v)$ of the *longest shortest path* (i.e., the longest graph geodesic) between any two graph vertices (u,v) of a graph, where $d(u,v)$ is a graph distance (in other words, graph diameter is the largest number of vertices which must be traversed in order to travel from one vertex to another when paths which backtrack, detour, or loop are excluded from consideration; Weisstein 2010); (iii) clustering coefficient: density associated with the open neighborhood of a given node (Watts and Strogatz 1998).

We introduce here a branching diagram or *cleavogram* that shows the splitting sequence of the different groups of taxa when NAM is performed. The cleavogram represents the arrangement of vertices into groups as the iterative removal of intermediary species progresses (Fig. 3a). Two kinds of branches can be distinguished in a cleavogram: divisible and indivisible. In turn, indivisible branches may be interrupted or complete (Fig. 3b). Taxa associated with interrupted branches represent intermediary species that are removed along the process. Complete branches contain entities not susceptible to further removal (isolated nodes, connected diads, and cliques) and remain unchanged until the last subnetwork. On the contrary, divisible branches consist of groups of taxa not fully connected and for this reason still susceptible to further cleavage. A pair of branches can be disjoint or compatible according to their species composition (Fig. 3b). The lists of species are mutually exclusive in disjoint branches, whereas in a pair of compatible branches, one is a subset of the other.

In the cleavogram, branches that meet the cohesiveness criteria are identified and will be candidate branches for network partitioning. The next step consists of a flat partition assigning species to independent groups, that is, selecting disjoint branches from the whole set of candidate branches. Under the contextual cohesiveness criteria, disjoint branches are chosen so that the partition is guaranteed to retrieve: (i) cohesive groups by definition; (ii) species directly or indirectly connected to another species belonging to the same group (groups are components); and (iii)

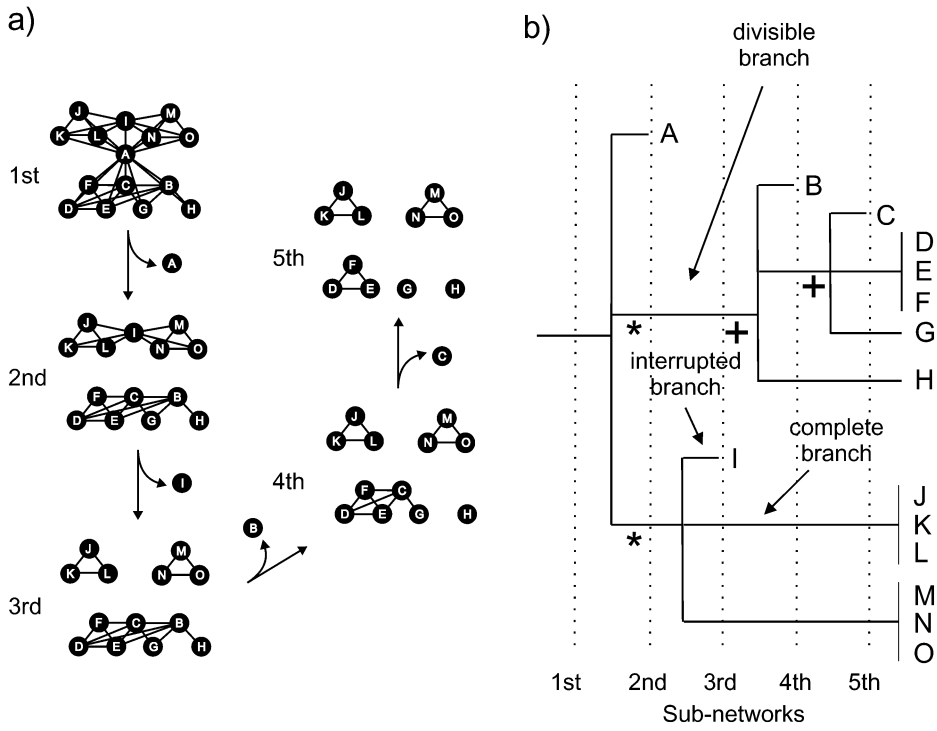

FIGURE 3. Graphical display of the sympatry network analysis. (a) Iterative removal of intermediary species (species with the highest betweenness score) and their incident edges. Process begins at the basal network and finishes when no further removal is possible. (b) Cleavogram associated with the network decomposition. The splitting sequence runs from the basal network (left of the cleavogram) to the last subnetwork (right of the cleavogram). Vertical dotted lines intersect the branches corresponding to components (set of connected species) of the respective subnetworks. Intermediary species are represented by interrupted branches as they are removed. Indivisible entities reach the terminal portion of the cleavogram. Examples of branch types from a morphological standpoint are indicated by arrows, whereas pairs of disjoint and compatible branches are marked with * and +, respectively.

no links running between groups (groups are disconnected unless there were intermediary species). Disjoint branches can be selected through three different procedures: (i) Forward search—first occurrence: the cleavogram is examined from the basal subnetwork to the terminal subnetwork, that is, from the most inclusive group to the most restrictive group, and candidate branches are selected as they are found; (ii) Forward search—last subdivision: the cleavogram is examined from the basal subnetwork to the terminal subnetwork, and a candidate branch is selected if it is no longer subdivided into mutually exclusive candidate branches, otherwise the search continues; (iii) Backward search: the cleavogram is examined from the terminal subnetwork to the basal subnetwork, and candidate branches are selected as they are found, discarding all the more basal (inclusive) ones. Given a universe of candidate branches, their set of involved species can be compatible or mutually exclusive. Forward search—first occurrence prefers candidate branches with the largest amount of elements from a series of compatible branches. This procedure should be preferred when connectivity is the priority in data sets with incomplete sampling (missing links expected). Forward search—last subdivision and Backward search choose the maximal set of mutually exclusive candidate branches. Backward search is the strictest procedure and should be preferred when the main concern is to recover the smallest areas supported by evidence. Forward search—last subdivision is concerned with branches of equal or larger cardinality to those selected by Backward search. It is a flexible heuristic tool that solves better the tradeoff between pattern resolution and supporting evidence. Figure 4 illustrates the resulting groupings from the application of the three different procedures.

The improvements on the inference and the analysis of the matrix described above are implemented in the software SyNet 2.0 available at http://www.cran.r-project.org, which is an add-on package for the statistical software R. Comparative features involving NAM as originally proposed and current improvements presented here are provided in Table 1.

Results

Empirical Example with Punctual Data

Ephemeroptera (mayflies) are an ancient lineage of insects, dating back to the late Carboniferous or early Permian periods (Barber-James et al. 2008). Aquatic nymphs are the dominant life history stage. Adults have short lives ranging from a few hours to a few days. Nymphs may disperse by drifting throughout river systems but are limited to those interconnected parts of the watershed above the influence of the sea. Successful dispersal by wind across natural barriers like the sea or mountain ranges is limited by the short life span of the adults. Therefore, the present distribution was believed to be a reflection of geological events (Edmunds 1972, 1975). Particularly, the mayfly fauna of New Zealand is a good example to illustrate our approach. All species are endemic to this archipelago, and their distribution patterns seem to be driven by landscape changes associated to Pleistocene glaciations (Hitchings 2008a). Additionally, mayflies are the dominant benthic invertebrates in the coldest upper reaches of glacial streams in New Zealand (Winterbourn et al. 2008). The worldwide Leptophlebiidae and the amphinotic Nesameletidae are the most prolific families in New Zealand with records of occurrences available from modern systematic works.

Data were obtained from specialized literature (Towns and Peters 1996; Hitchings and Staniczek 2003; Hitchings 2008b, 2009a; Winterbourn 2009). Localities given in New Zealand Map Grid projection metric coordinates were converted to the WGS84 datum with the online utility provided by the Land Information New Zealand government department (http://www.linz.govt.nz/). Additional records for *Zephlebia pirongia* were inferred from its dot map in Hitchings (2008b). Random samples without replacement of 30% of records previously extracted for *Z. pirongia* were assigned to its cogeneric homopatric species (*Z. dentata, Z. nebulosa, Z. inconspicua, Z. versicolor*) (Hitchings 2008a). Duplication of records was avoided to assess better the heuristic capabilities of our

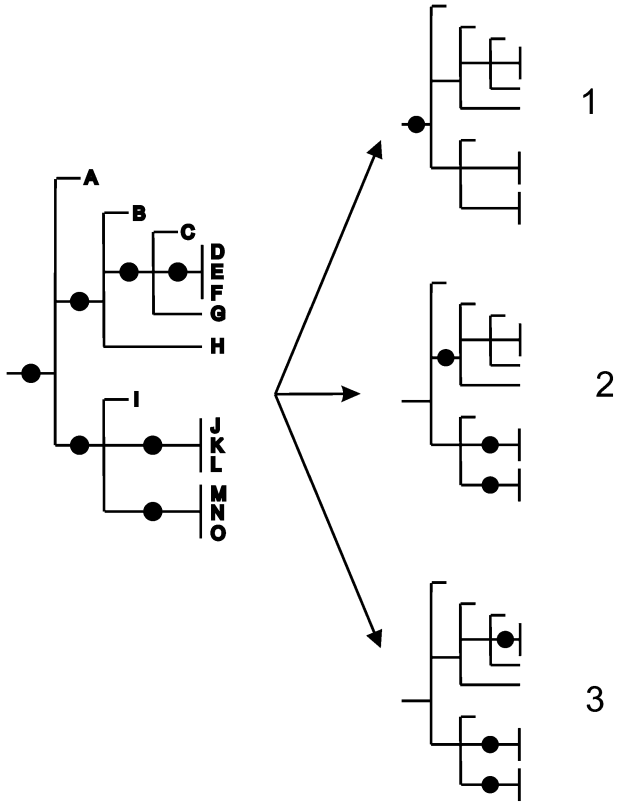

FIGURE 4. Network partitioning with different procedures. Same cleavogram as in Figure 3. Candidate branches (adjusting to a cohesiveness framework, e.g., graph diameter not higher than two) are marked with black dots. Selected branches differ according to procedure implemented. 1, Forward search—first occurrence; 2, Forward search—last subdivision; 3, Backward search.

TABLE 1. Comparison between original and modified NAM approaches

		NAM (original)	NAM (modified)
Inference instance	Input	Punctual records	Punctual records
	Criteria	• Interpenetration • Relative proximity	• Interpenetration • Relative proximity
	Procedure	Trimming of delaunay triangulation—radial screening for overlap	**Topological resemblance (MST) cost for homopatry (ACSH)**
	Output	Binary matrix	**Weighted matrix**
Analysis	Input	Binary basal network without discriminating strength of sympatry links	**Dichotomization →binary basal network discriminating strength of sympatry links**
	Criterion	Partition of the network into entities satisfying within-group sympatry and between-group allopatry	Partition of the network into entities satisfying within-group sympatry and between-group allopatry or *minor sympatry*
	Procedure	• Betweenness score—iterative removal • Overall clustering performance for the whole network • Selection of components from one subnetwork selected	• Betweenness score—iterative removal • **Cohesiveness metrics restricted to components** • **Selection of components from one or more subnetworks**
	Output	• Network partitions (UC)	• Network partitions (UC) • *Cleavogram*
Spatial expressions	Criterion	Union of species records belonging to the same UC	Union of species records belonging to the same UC
	Output	Dot maps	Dot maps

Notes: Rows show similarities and differences at equivalent levels. Columns summarize sequence of steps in each version. Changes in bold and extensions in italics.

approach. The file containing information on each of the 679 compiled records is available from the online Appendix 1 (available from http://www.sysbio.oxfordjournals.org/) including species label, geographical coordinates in decimal format, and scientific source.

The original and reweighted matrices of sympatric association and the table of ACSH scores are available in Tables S1–S3. A simple linear normalization to map the reweighted values onto the interval between 0 and 1 was performed. The basal network was obtained from reweighted topological resemblance >0.8 and ACSH <100 km. These thresholds allow the separation of approximately 80% of strongest pairwise relationships between species given by stable marriage coupling (Gale and Shapley 1962) found in the data set. The cleavogram derived through NAM analysis on the basal network is shown in Figure 5a. Branches including three or more species and density cohesion ≥ 0.1 were considered candidates. After applying the above methodology, NAM recovers four and six UCs with Forward search—first occurrence and Backward search, respectively, that are comparable to the patterns proposed by Hitchings (2008a). The most distinctive groups are shown: (i) Northern North Island (Fig. 5b); (ii) Eastern South Island southward of Alpine Fault (Southern Alps) (Fig. 5c); (iii) Widely distributed on both islands (Fig. 5d); iv) North Island + NW South Island (Fig. 5e). Some restricted distributions were identified as isolated nodes from the onset of the analysis: Auckland Island, Banks Peninsula, and Northernmost tip of North Island.

New Zealand biota distribution has been shaped by a complex suite of historical factors: Gondwanic footprints, volcanism, Pleistocene glaciation, alpine orogenesis, sea level oscillations, colonization, and extinction processes associated to oceanic archipelago (Goldberg et al. 2008). The analysis on the Ephemeroptera data recovers the signature of some of these influential factors. Following, the most relevant agreements between patterns and hypothetical historical drivers are pointed out. Group 1 (formed by three species of different genera) matches the Northern North Island province proposed by Leathwick et al. (2007) in their classification of freshwater ecosystems. They consider that this province is typified by relatively muted impacts during the Last Glacial Maximum. Accordingly, the respective forest cover remained dominant resulting in a much more stable landscape than occurred over much of the rest of New Zealand at that time. McGlone (1985) also recognized this region in his plant biogeography study, where the high endemism level of trees in this geographical unit was stressed. Group 2 includes three species of *Deleatidium* mainly confined to cold glacier-fed lotic freshwaters in the Southern Alps (Hitchings 2009b). The pooled distribution of these species is concordant with the diverse and extensive alpine biota (e.g., Gibbs 2006). According to McGlone (1985), the rapidly rising Southern Alps may have acted as a center of speciation because of its provision of novel alpine and subalpine environments. Craw et al. (2008) claim that biological dispersal across the Southern Alps may have been facilitated by numerous mountain passes, especially via the new passes formed by crosscutting faults. Noticeably, the mountain belt of the Southern Alps (developed in the NE–SW direction parallel to the Alpine

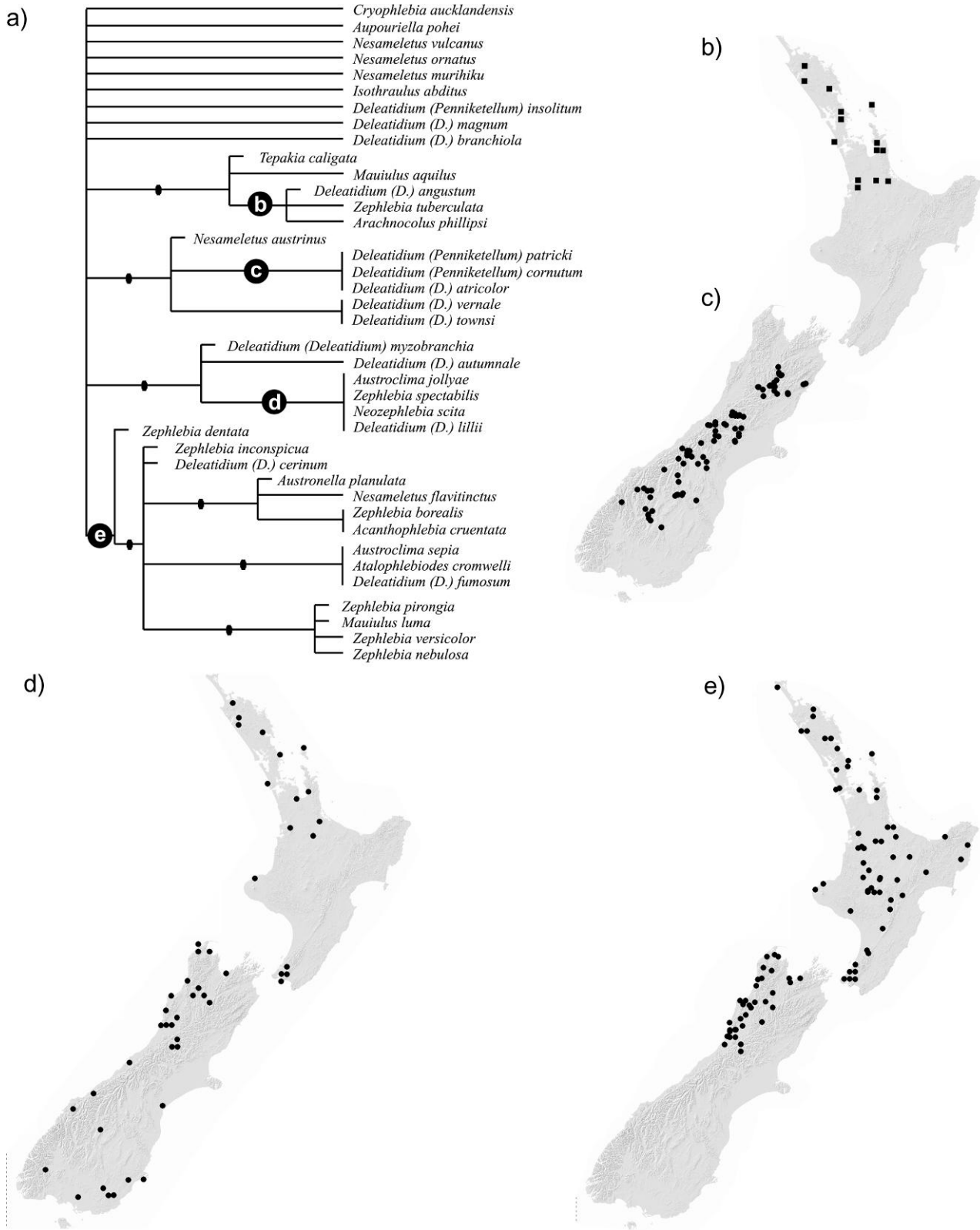

FIGURE 5. Empirical example 1: New Zealand Ephemeroptera. (a) Cleavogram showing candidate (black dots) and selected branches (lettered circles). (b–e) Spatial expressions of UCs labeled as in cleavogram. (b) Northern North island. (c) Eastern South Island. (d) Widely distributed on both Island. (e) North Island + NW South Island. Shaded relief map obtained from the digital mapping company Geographx (http://www.geographx.co.nz).

Fault) has apparently acted as an effective barrier with regards to this group. Group 3 is composed by four species of different genera and it is widespread on both islands. A remarkable feature for this flexible group is the absence of records from the Central and Southern North Island Provinces (sensu Leathwick et al. 2007) with the exception of Wellington and Taranaki areas. Apparently, this pattern would reflect the disturbance promoted by volcanism and the defaunation associated to the discharged volcanic material in this region (McDowall 1996). Group 4 (North Island + NW South Island) would reflect an ancient distribution spanning over what used to be a continuous unglaciated area presently represented by disjoint land masses. According to Hitchings (2008a), this striking pattern can be considered a consequence of the combination of lowered sea levels and the ice sheet barrier confining populations during Pleistocene glaciation to the underlying area of this distinctive mayfly assemblage.

Empirical Example with Grids

Goldani et al. (2006) coded the distribution of 106 Neotropical primate (Platyrrhini) species into 60 quadrats of 5° latitude by 5° longitude as OGUs, detecting eight areas of endemism (Fig. 6a). We used this data matrix as input to apply the weighted approach and infer the sympatry network. Indices for rows and columns were used as arbitrary coordinates for individual records (Fig. 6b). The file with the complete list of species and respective coordinates of occurrence is available on the online Appendix 2. Entries of the resulting matrix of ACSH spatial affinity with values $\leq$0.5 were coded as 1 in the sympatric matrix and otherwise coded as 0. Approximately 80% of strongest pairwise relationships in the stable marriage coupling fall below this threshold. This is also strongly correlated with a scenario of co-occurrence in at least one cell and spatial contiguity of occupied cells. The respective network was analyzed with NAM yielding the cleavogram shown in Figure 7a. Candidate branches with densities $\geq$0.1 are marked. Two different network partitioning strategies were carried out in this cleavogram: (i) Forward search—first occurrence and (ii) Forward search—last subdivision. The first strategy yielded three main UCs highlighted with numbered circles that broadly match the Chacoan + Parana, Caribbean, and Amazon Basin subregions (Morrone 2001; Nihei and De Carvalho 2007), although there was minor overlap along their borders. The second strategy, with higher resolution, produced subdivisions (lettered stars) of the previously obtained UCs, which represent finer details associated to groups of tightly codistributed species. In this way, it is possible to refer the patterns to relevant ecoregions (Olson et al. 2001). For example, Caatinga and Atlantic Forests can be recognized within the Chacoan + Parana group, whereas Isthmian and Chocó-Darién Moist Forests are inside the Caribbean entity. Differently from the four areas detected by Goldani et al. (2006), the present study recognizes eight UCs within the Amazon region. Some spatial expressions spread along the main fluvial axis of the Amazon and tributaries. Therefore, the cleavogram allows here the arrangement of species distributions into a nested sequence of increasingly restricted distributions.

Hypothetical Examples

Casagranda et al. (2009) presented six hypothetical examples (plus two variations) to show situations, in which NAM, as proposed in 2008, would fail to handle sympatric patterns (for details, see their figures and descriptions of cases). In its original form, NAM operated on basal networks without discriminating the strength of sympatry links, thus the resolution of patterns obtained was reduced. Problems in the inference instance did not allow the recognition of certain patterns in the examples provided. These hypothetical examples were reanalyzed (except the one corresponding to their Fig. 2c, not available in their supplementary material), with the improvements presented in this paper. NAM found all the expected patterns from all these examples using the default search provided by SyNet 2.0. Furthermore, NAM found finer patterns composed of five clusters based on the coextensive sympatry that apparently passed unnoticed to their scrutiny in their own Figure

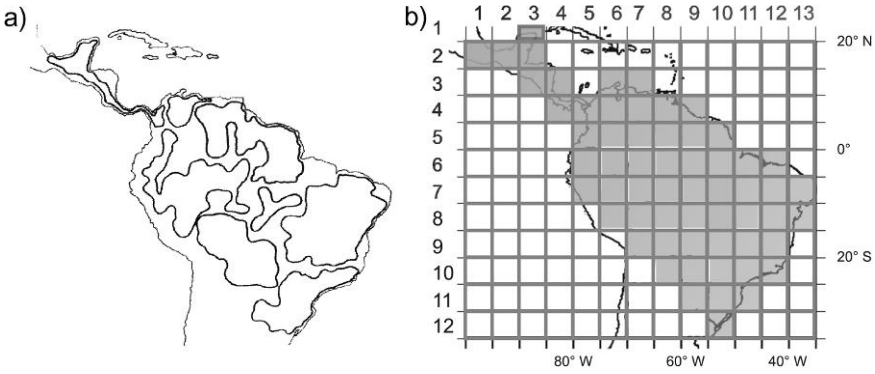

FIGURE 6. (a) Areas of endemism based on Neotropical primates according to Goldani et al. (2006). (b) Grid reference. Indices of rows and columns define the coordinates for each cell record.

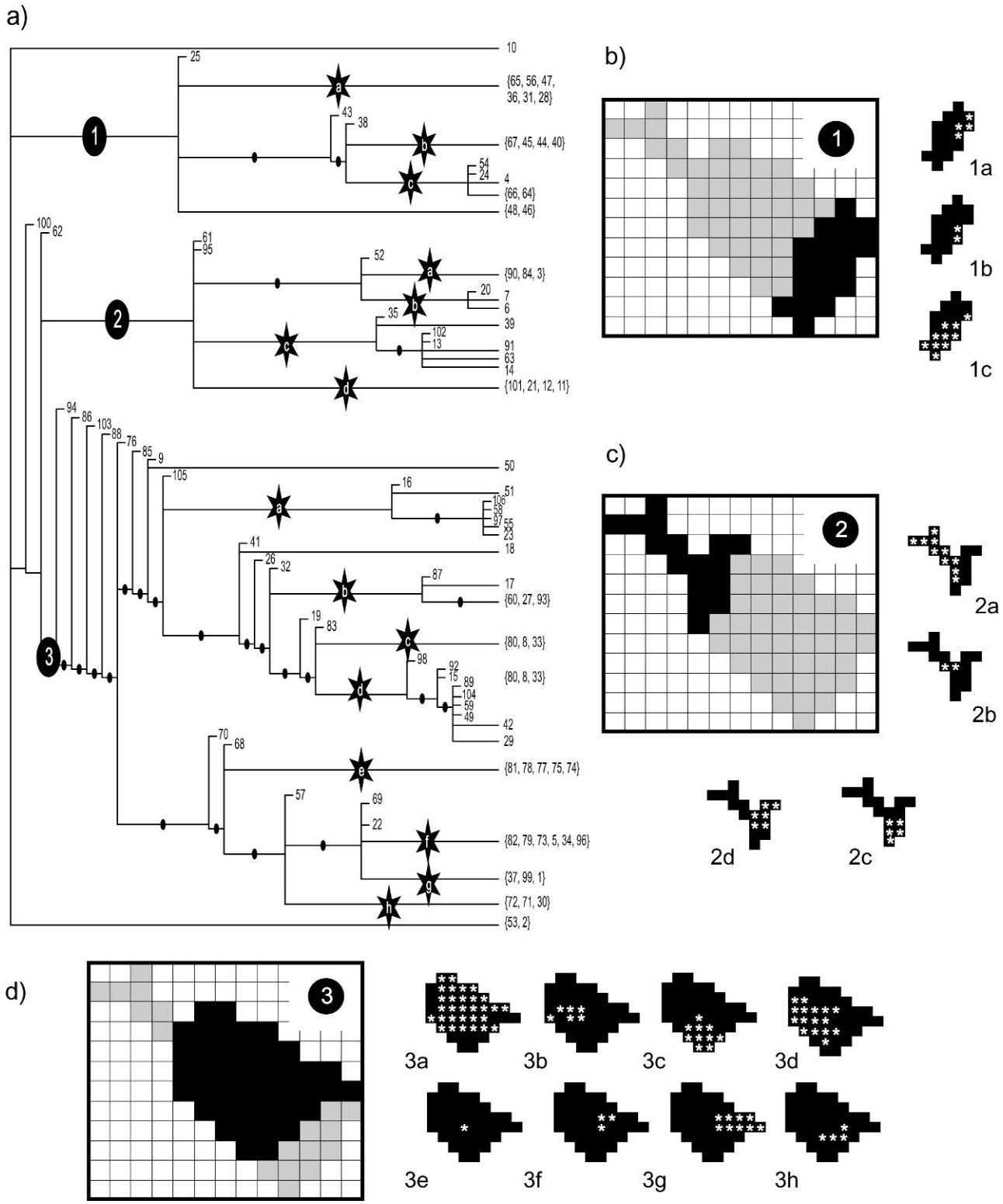

FIGURE 7. Empirical example 2: Neotropical primates (Platyrrhini). (a) Cleavogram showing sequence of group separation through intermediary species removal. Numbers at terminals represent species labels following table 1 of Goldani et al. (2006). Branches complying with the density criterion $\geq$0.1 are marked. Black circles with numbers show the three major UCs (1 = Chacoan + Parana, 2 = Caribbean, and 3 = Amazon Basin) according to Forward search—first occurrence. Stars with letters indicate the further subdivision of these UCs, into subgroups of codistributed species identified by Forward search—last subdivision. (b–d) Spatial expressions of UCs.

1b. Resolution of matrices and spatial expressions of UCs found based on these examples are provided in online Appendix 3).

Discussion and Conclusions

Intricate distributions of species result in a reticulated structure of spatial associations. The study of these sympatric relationships within a theoretical network framework seems appropriate because sympatry is a relational datum in itself. NAM seeks groups of species embedded into the global network that are internally sustained by sympatric cohesiveness but weakly connected (or disconnected) to outgroup entities. Several improvements to the original NAM approach are introduced in this paper:

1) Development of measures sensitive to different degrees of range overlap, appropriate for the construction of weighted networks of spatial association based on set of punctual records.

2) Dynamic analysis of the network: identification of components in the different subnetworks fulfilling the criteria of cohesiveness.

3) Design of a graphic tool (cleavogram) that summarizes the structure of the network revealed by the iterative removal of the intermediary species.

Weighted relationships help to discriminate coextensive sympatry from minor overlap between species ranges. This is an important issue related to the identification of supporting elements for candidates to areas of endemism. This approach addresses what Dos Santos et al. (2008, p. 446) already advanced and enables NAM to overcome the problems highlighted by Casagranda et al. (2009). The topological resemblance was evaluated using the MST as a basic tool of study. The MST is a curvilinear object but it is still linked to the idea of an underlying area from where records have been drawn. Accordingly, Carmi et al. (2005) have proven that for a given set of points in the plane, the MST is a constant-factor approximation for the minimum-area spanning tree problem. The use of MST is not new to biogeographic studies since it has been applied in the context of quantitative panbiogeography (e.g., Page 1987; Craw et al. 1999). The developed measure allows the recognition of the degree of association among complex dot clouds and the distinction of overlapped distributional patterns. The statistical significance of the coefficients of sympatric association is still an open field of research. One possibility to address this subject would be to compare observed with randomly generated values via permutation of dot sets, but this option has not been explored yet.

Several authors have criticized the use of hard thresholds to dichotomize a weighted network (e.g., Hanneman and Riddle 2005; Butts 2009; Opsahl and Panzarasa 2009). Their main reasons are loss of information and arbitrary choice of cutoff values. On the other hand, the extraction of meaningful information of weighted networks has been considered a challenging task, and proper generalization of network measurements obtained from unweighted graphs (e.g., betweenness, clustering coefficient) is a subject much debated (Ahnert et al. 2007; Opsahl et al. 2010; Abdallah 2011). As NAM focuses on the identification of intermediacy in the information flow through the network, it is desirable that the betweenness score is calculated in function of paths composed of edges with comparable weights. Dichotomization appears as the best alternative at the moment, as it recovers strong links under the appropriate threshold choice and produces a workable and interpretable matrix. We dichotomize the weighted network so that 80% of values found on the stable marriage coupling are above the selected threshold, empirically determined to recover the main structure of the strongest links. The input binary matrix used for NAM in Dos Santos et al. (2008) differs from the one used here. In the original approach, the binary matrix was a simplistic arrangement of relations undisturbed by their relative position in a graded series of strengths. Zero entries of the matrix meant allopatry. Now, the binary matrix represents a collection of links deemed to be meaningful after considering their weights, and zero entries of the matrix mean either allopatry or weaker sympatric association between elements.

The cleavogram illustrates the spatial relationships between species within a network context. It is a simplified way to show the splitting sequence of groups as the removal of intermediary elements proceeds. It was not devised to depict a hierarchical clustering of species in function of their distributions. If the cleavogram exhibits many successive interrupted branches, it may be indicating that unseen patterns are associated to the pool of intermediary elements. Eventually, a residual analysis applied to those elements may recover additional patterns. However, we have empirically observed that increasing the level of strictness for considering sympatry may extract additional groups otherwise merged into a comb of interrupted branches.

The original approach selected one single subnetwork, whereas the current approach identifies UCs based on their intrinsic cohesiveness. On the contrary, the results of the new proposal unmask patterns lying on different subnetworks generated by the removal process. This can be visualized in the cleavogram with branches selected at different levels. It is noteworthy that another search algorithm for grouping based on different dendrogram levels has been recently developed (Gurrutxaga et al. 2010).

When incorporating weights and group cohesiveness into the analysis of the sympatry network, our approach becomes a flexible tool to address many different queries on pattern search. If we define sympatry by the mere overlap of species ranges (disregarding the degree of overlap), the resulting partition will render allopatric groups that eventually may show internally structural holes. Additionally, if cohesiveness is considered under the most stringent setting (selected groups must be cliques), the resulting partition will satisfy the duality within-group sympatry and between-group allopatry across the set of species considered. Finally, if we

equate cohesiveness with clique status and differentiate major from minor overlap, our analysis will recover groups of species fully connected by coextensive sympatry. In addition, intragroup links will be stronger than intergroup ones. Spatial expressions of the resulting groups will constitute an assemblage of areas not necessarily disjoint from each other but certainly supported by pools of two or more unique, endemic, and codistributed species. Groups not perfectly cohesive are also of interest to our approach because they can reveal meaningful geographical information that would be lost if only a scenario of supporting species identically distributed is expected.

After the pattern resolution is obtained by NAM, the next step is to establish whether or not historical processes have driven the structure of the network. Network theory provides a good platform for an integrative approach, allowing the study of the sympatry network with phylogenetic metadata. For example, if the network structure was induced by vicariance, then closely related species should be disconnected in the network appearing in disjoint branches of the cleavogram. This could be studied with the network concepts, like heterophily and assortative mixing (Newman 2003; Park and Barabási 2007). A preliminary advance on this topic was presented at the VI Southern Connection Congress, Bariloche, Argentina (Molineri et al. 2010). In conclusion, as in original NAM, spatial expressions of UCs are building blocks for biogeography supported by strictly endemic and connected species through sympatry paths.

Supplementary Material

Supplementary material, including data files and/or online-only appendices, can be found at http://www.sysbio.oxfordjournals.org/.

Funding

This work was supported by the Argentinean National Council of Scientific and Technological Research (CONICET) (PIP 1424) and the Argentinean National Agency for the Promotion of Science and Technology (PICT 528).

Acknowledgments

We would like to thank A. S. H. Breure, H. R. Fernández, and G. Wibmer for their suggestions to an early draft of the manuscript. R. Deutsch kindly read the final draft of this manuscript making suggestions on both English and sympatry strength measures. Editors R. DeBry and A. Paterson contributed to strengthen the manuscript, and the anonymous reviewers provided valuable criticism and constructive suggestions. E. Santos Discépolo contributed to a comfortable workplace and environment. Authors work for CONICET, whose support is greatly acknowledged.

References

Abdallah S. Forthcoming 2011. A new methodology for generalizing unweighted network measures. Social network analysis and mining. Springer.

Ahnert S.E., Garlaschelli D., Fink T.M.A., Caldarelli G. 2007. Ensemble approach to the analysis of weighted networks. Phys. Rev. E. 76:016101.

Barber-James H.M., Gattolliat J.L., Sartori M., Hubbard M.D. 2008. Global diversity of mayflies (Ephemeroptera, Insecta) in freshwater. Hydrobiologia. 595:339–350.

Barrat A., Barthélemy M., Pastor-Satorras R., Vespignani A. 2004. The architecture of complex weighted networks. Proc. Natl. Acad. Sci. U.S.A. 101(11):3747–3752.

Brandes U., Erlebach T., editors. 2005. Network analysis: methodological foundations. LNCS 3418. Berlin-Heidelberg (Germany): Springer.

Butts C.T. 2009. Revisiting the foundations of network analysis. Science. 325(5939):414–416.

Carmi P., Katz M.J., Mitchell J.S.B. 2005. The minimum area spanning tree problem. Proceedings Workshop on Algorithms and Data Structures (WADS), Waterloo (Canada): Springer LNCS 3608. p. 195–204.

Casagranda D., Arias J.S., Goloboff P.A., Szumik C., Taher L.M., Escalante T., Morrone J.J. 2009. Proximity, interpenetration, and sympatry networks: a reply to Dos Santos et al. Syst. Biol. 58: 271–276.

Craw D., Burridge C.P., Upton P., Rowe D.L., Waters J.M. 2008. Evolution of biological dispersal corridors through a tectonically active mountain range in New Zealand. J. Biogeogr. 35: 1790–1802.

Craw R.C., Grehan J.R., Heads M.J. 1999. Panbiogeography. Tracking the history of life. New York (NY): Oxford University Press.

Dolan J., Weiss R., MacGregor Smith J. 1991. Minimal length tree networks on the unit sphere. Ann. Oper. Res. 33:503–535.

Dos Santos D.A., Deutsch R. 2010. The positive matching index: a new similarity measure with optimal characteristics. Pattern Recognit. Lett. 31:1570–1576.

Dos Santos D.A., Fernández H.R., Cuezzo M.G., Domínguez E. 2008. Sympatry inference and network analysis in biogeography. Syst. Biol. 57:432–448.

Edmunds G.F. 1972. Biogeography and evolution of Ephemeroptera. Annu. Rev. Entomol. 17:21–42.

Edmunds G.F. 1975. Phylogenetic biogeography of mayflies. Ann. MO. Bot. Gard. 62:251–263.

Freeman L.C. 1977. A set of measures of centrality based on betweenness. Sociometry. 40:35–41.

Gale D., Shapley L. 1962. College admissions and the stability of marriage. Am. Math. Mon. 69:9–15.

Gibbs G. 2006. Ghosts of Gondwana—the history of life in New Zealand. Nelson (New Zealand): Craig Potton.

Goldani A., Carvalho G.S., Bicca-Marques J.C. 2006. Distribution patterns of Neotropical primates (Platyrrhini) based on parsimony analysis of endemicity. Braz. J. Biol. 66(1a):61–74.

Goldberg J., Trewick S.A., Paterson A.M. 2008. Evolution of New Zealand's terrestrial fauna: a review of molecular evidence. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363:3319–3334.

Gurrutxaga I., Albisua I., Arbelaitz O., Martín J.I., Muguerza J., Pérez J.M., Perona I. 2010. SEP/COP: an efficient method to find the best partition in hierarchical clustering based on a new cluster validity index. Pattern Recognit. 43(10):3364–3373.

Hanneman R.A., Riddle M. 2005. Introduction to social network methods. Riverside (CA): University of California, Riverside. Available from: http://faculty.ucr.edu/~hanneman/.

Hitchings T.R. 2008a. The post glacial distribution of New Zealand mayflies. In: Hauer F.R., Stanford J.A., Newell R.L., editors. International advances in the ecology, zoogeography and systematics of mayflies and stoneflies. Volume 128. Berkeley (CA): University of California Publications in Entomology. p. 89–101.

Hitchings T.R. 2008b. A new species of *Deleatidium (Penniketellum)* and the adult of *D. (P.) cornutum* Towns and Peters (Ephemeroptera: Leptophlebiidae) from New Zealand. Rec. Canterbury Museum. 22:31–43.

Hitchings T.R. 2009a. Three new species of *Deleatidium* (*Deleatidium*) (Ephemeroptera: Leptophlebiidae) from New Zealand. Rec. Canterbury Museum. 23:35–50.
Hitchings T.R. 2009b. Leptophlebiidae (Ephemeroptera) of the alpine region of the Southern Alps, New Zealand. Aquat. Insects. 31(1 Suppl):595–601.
Hitchings T.R., Staniczek A.H. 2003. Nesameletidae (Insecta: Ephemeroptera). Fauna of New Zealand 46. Canterbury, Lincoln (New Zealand): Manaaki Whenua Press. p. 1–72.
Humphries C.J., Parenti L. 1999. Cladistic biogeography: interpreting patterns of plant and animal distributions. New York: Oxford University Press.
Leathwick J.R., Collier K., Chadderton L. 2007. Identifying freshwater ecosystems with nationally important natural heritage values: development of a biogeographic framework. Science for conservation 274. Wellington (New Zealand): Science & Technical Publishing.
Legendre P., Legendre L. 1998. Numerical ecology. Amsterdam (The Netherlands): Elsevier Science BV.
Mast A., Nyffeler R. 2003. Using a null model to recognize significant co-occurrence prior to identifying candidate areas of endemism. Syst. Biol. 52:271–280.
McDowall R.M. 1996. Volcanism and freshwater fish biogeography in the northeastern North Island of New Zealand. J. Biogeogr. 23:139–148.
McGlone M.S. 1985. Plant biogeography and the late Cenozoic history of New Zealand. N.Z. J. Bot. 23:723–749.
Molineri C., Nieto C., Domínguez E., Dos Santos D.A. 2010. Biogeography of South American Ephemeroptera. Austral insect patterns: phylogenetics and biogeography of austral insects. VI Southern Connection Congress, San Carlos de Bariloche, Río Negro (Argentina): Universidad Nacional del Comahue.
Morrone J.J. 2001. Biogeografía de América Latina y el Caribe. Manuales y Tesis SEA, 3. Zaragoza (Spain): Sociedad Entomológica Aragonesa.
Newman M.E.J. 2003. Mixing patterns in networks. Phys. Rev. E. 67(2):026126.
Newman M.E.J. 2010. Networks: an introduction. Oxford (UK): Oxford University Press.
Nihei S.S., De Carvalho C.J.B. 2007. Systematics and biogeography of *Polietina* Schnabl & Dziedzicki (Diptera: Muscidae): Neotropical areas relationships and Amazonia as composite area. Syst. Entomol. 32:477–501.
Olson D.M., Dinerstein E., Wikramanayake E.D., Burgess N.D., Powell G.V.N., Underwood E.C., D'amico J.A., Itoua I., Strand H.E., Morrison J.C., Loucks C.J., Allnutt T.F., Ricketts T.H., Kura Y., Lamoreux J.F., Wettengel W.W., Hedao P., Kassem K.R. 2001. Terrestrial ecoregions of the world: a new map of life on earth. BioScience. 51:933–938.
Opsahl T., Agneessens F., Skvoretz J. 2010. Node centrality in weighted networks: generalizing degree and shortest paths. Soc. Networks. 32:245–251.
Opsahl T., Panzarasa P. 2009. Clustering in weighted networks. Soc. Networks. 31:155–163.
Page R.D.M. 1987. Graphs and generalized tracks: quantifying Croizat's panbiogeography. Syst. Zool. 36:1–17.
Parenti L.R., Ebach M.C. 2009. Comparative biogeography: discovering and classifying biogeographical patterns of a dynamic earth. Berkeley (CA): University of California Press.
Park J., Barabási A.L. 2007. Distribution of node characteristics in complex networks. Proc. Natl. Acad. Sci. U.S.A. 104:17916–17920.
Prim R.C. 1957. Shortest connection networks and some generalizations. Bell Syst. Tech. J. 36:1389–1401.
Towns D.R., Peters W.L. 1996. Leptophlebiidae (Insecta: Ephemeroptera). Fauna of New Zealand 36. Canterbury, Lincoln (New Zealand): Manaaki Whenua Press. p. 1–141.
Wasserman S., Faust K. 1999. Social network analysis: methods and applications. Structural analysis in the social sciences 8. New York (NY): Cambridge University Press.
Watts D.J., Strogatz S.H. 1998. Collective dynamics of "small-world" networks. Nature. 393:440–442.
Weisstein E.W. 2010. Graph diameter. From *MathWorld*—a Wolfram web resource. Available from: http://mathworld.wolfram.com/GraphDiameter.html.
Winterbourn M.J. 2009. A new genus and species of Leptophlebiidae (Ephemeroptera) from northern New Zealand. N.Z. J. Zool. 36:423–430.
Winterbourn M.J., Cadbury S., Ilg C., Milner A. 2008. Mayfly production in a New Zealand glacial stream and the potential effect of climate change. Hydrobiologia. 603:211–219.