On the use of median-joining networks in evolutionary biology

Sungsik Kong[a,†], Santiago J. Sánchez-Pacheco[a,†] and Robert W. Murphy[a,b,*]

[a]*Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, ON M5S 3B2, Canada;* [b]*Department of Natural History, Royal Ontario Museum, 100 Queen's Park, Toronto, ON M5S 2C6, Canada*

Abstract

Median-joining (MJ) was proposed as a method for phylogeographical analysis and is enjoying increasing popularity. Herein, we evaluate the efficacy of the approach as originally intended. We show that median-joining networks (MJNs) are theoretically untenable for evolutionary inference, and that confusion has afflicted their use for over 15 years. The approach has two obvious shortcomings: its reliance on distance-based phenetics (overall similarity instead of character transformations) and the lack of rooting (no direction or history). Given that evolution involves both change and time, and the absence of rooting removes time (ancestor–descendant relationships) from the equation, the approach cannot yield defensible evolutionary interpretations. We also examine the impact of MJ analyses on evolutionary biology via an analysis of citations and conclude that the spread of MJNs through the literature is difficult to explain, especially given the availability of character-based analyses.
© The Willi Hennig Society 2015.

The rising preeminence of Phylogenetic Systematics runs the risk of being self defeating, for it is becoming more and more common for practitioners of other approaches to pay lip-service to phylogenetic principles …This tendency seems to be most pronounced when the alternative approaches are of a mathematical nature or are implemented by computer programs, and the practice hinders continued development of truly phylogenetic methods. [Farris et al., 1982, p. 317]

Technological advances in computing and the flood of molecular data have catapulted research in evolutionary biology in general, and in phylogenetics in particular, by testing hypotheses through user-friendly software. However, these advances have also prompted the proliferation of sophisticated-looking analyses without any consideration of the philosophy behind the methods (Grant et al., 2003). In this context, the median-joining (MJ) approach has been implemented for over 15 years. The application of median-joining networks (MJNs) to evolutionary studies has grown dramatically, and there is no indication that this trend

will wane in the foreseeable future (Fig. 1). As is necessary for many widely accepted approaches and concepts, such as the biological species concept, it is important to re-evaluate the assumptions and limitations of MJ. Herein, we review its theoretical foundations, current applications and the associated terminology. We discover that its implementation has been plagued by confusion since its conception. The approach overlooks basic principles of both evolutionary biology and phylogenetic analysis, and even the underlying prerequisites of MJ itself.

Bandelt et al. (1999) (hereinafter BEA99) introduced MJ as a method for inferring intraspecific phylogenies, stating that "[r]econstructing *phylogenies* from intraspecific data…is often a challenging task because of large sample sizes and small genetic *distances* between individuals" (emphasis added). They argued further that "[t]he resulting multitude of plausible trees is best expressed by a *network* which displays alternative potential *evolutionary paths in the form of cycles*" [p. 37 (abstract), emphasis added]. MJ constructs such networks. Although the non-evolutionary essence of MJ analysis was summarized exceptionally well in the two consecutive, introductory sentences, MJNs

*Corresponding author:
E-mail address*: bob.murphy@utoronto.ca
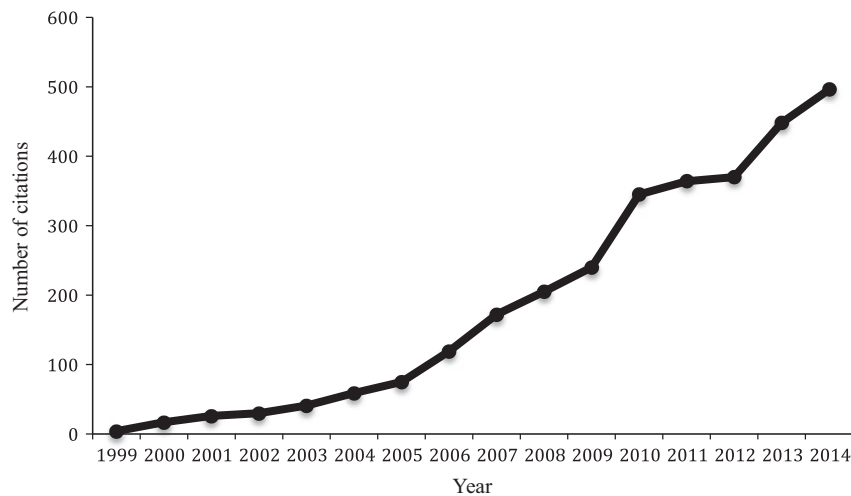[†]S.K. and S.J.S-P. contributed equally.

Fig. 1. Absolute number of citations of Bandelt et al. (1999) between 1999 and (December) 2014 (ISI Web of Science).

continue to persist and flourish in the literature. Therefore, it is desirable to address the most conspicuous drawbacks of this method: its distance-based approach and the absence of rooting.

Theoretical overview

Phylogenetic networks: rooting matters

A phylogenetic tree is a directed (rooted) branching diagram that represents the hypothesized relationships among the organisms under study. Unlike undirected (unrooted) branching diagrams (i.e. networks; but see below), historical statements must be made based on trees (Wheeler et al., 2006) because outgroup comparison roots the ingroup topology and polarizes character transformations (Farris, 1972, 1982), thereby converting a non-evolutionary network into an evolutionary hypothesis. Although both trees and networks are cladograms (Farris, 1970)—diagrams with a branching pattern—that depict hypothesized relationships, the evolutionary history of the evidence (e.g. DNA sequences) and the organisms can only be inferred by ordering terminals, explaining characters and testing hypotheses on trees. However, because of the bifurcating pattern of phylogenetic trees (two, and only two, descendant branches arise from a single ancestral branch), reticulation events derived from non-vertical inheritance processes such as hybridization, recombination and horizontal gene transfer cannot be visualized. Alternatively, networks are used frequently to represent such events.

The concept of a phylogenetic network has been used indiscriminately in the literature. In mathematics and computer science (specifically, in graph theory), a "network" is an undirected cyclical graph (UCG), which is nothing more than an unrooted branching diagram with reticulation. This differs from a "tree", which is a directed acyclical graph (DAG), or a rooted branching diagram without reticulation (Wheeler et al., 2006; Wheeler, 2012). In UCGs, a cycle is formed by a path over edges from a vertex back to itself where each intermediate edge between vertices is visited once. In contrast, cycles cannot be formed in DAGs because there is only one unique path between two different vertices, i.e. the edges between vertices can be traversed in one direction only. For vertex v in graph G, the *degree* is the number of edges in G that contains v. In DAGs, in-degree and out-degree edges can be specified, and their sum is the degree. For instance, in Fig. 2 the degree of vertex v_4 is 3, its in-degree and out-degree being 1 and 2, respectively. Note that trees are connected graphs because there are no vertices with degree 0 (all vertices are visited by a path over edges). They are composed of three types of vertices: the root (in-degree 0 and out-degree 2), the internal vertices (in-degree 1 and out-degree 2) and the leaves or terminals (in-degree 1 and out-degree 0) (Chung, 1986; Moret et al., 2004; Wheeler, 2012, 2014).

In a phylogenetic context, trees are basically a series of ancestor–descendant statements, as well as representations of sister-group relationships. Accordingly, the nodes (= vertices) signify both sister groups and ancestral conditions, and the branches (= edges) that connect them contain the character transformations between ancestors and descendants (Nelson, cited in Eldredge and Cracraft, 1980; Wheeler, 2012). Considering that "phylogenetics" refers to the evolutionary history of sets of organisms, and that direction through rooting is imperative to allow evolutionary inference, a "phylogenetic network" is then defined as a DAG with at least one node with in-degree 2 and
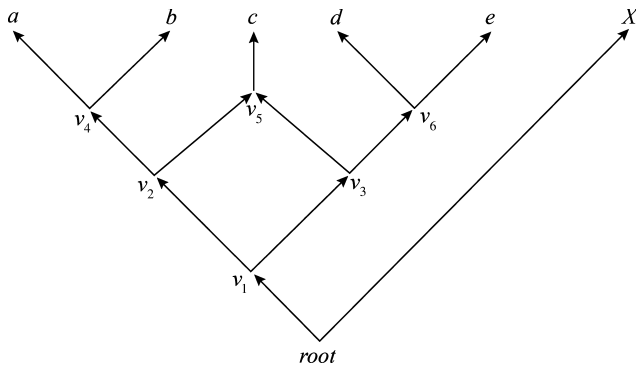
Fig. 2. A phylogenetic network. Note the absence of cycles due to the direction of the edges (arrows). A reticulation event is represented by a node with in-degree 2 and out-degree 1 (v_5).

out-degree 1 (Moret et al., 2004; Wheeler, 2012; Fig. 2: v_5), which denotes a reticulation event. In other words, a phylogenetic network is a tree with directed reticulate branches (Wheeler, 2014). By comparison, an MJN is an UCG[1] (Fig. 3; "cyclical evolution" occurs) or an unrooted[2] branching diagram with reticulation (= network only), and is therefore not a phylogenetic network as introduced by BEA99 and as often understood. Considering these criteria, the statistical parsimony network (Templeton et al., 1992), among other UCGs, is also a non-evolutionary network. Conversely, hybridization (Maddison, 1997; Linder and Rieseberg, 2004), recombination (Griffiths and Marjoram, 1996; Song and Hein, 2005) and duplication-loss-transfer (DLT) (Delwiche and Palmer, 1996; Planet et al., 2003) networks can be considered as phylogenetic, evolutionary networks. Phyletic group-types for phylogenetic networks, in addition to the Hennigian mono-, para- and polyphyletic, are defined by Wheeler (2014).

The (distance-based) MJ method

Operational details of the MJN algorithm are available (BEA99; Huson et al., 2010). Here, we focus on its phenetic nature. Unlike its preceding reduced med-

---

[1] Sometimes MJ analysis results in an undirected *acyclical* graph (i.e. an undirected, simple graph; e.g. Gangloff et al., 2013).

[2] Although the program NETWORK (fluxus-engineering.com; BEA99) offers the option to "root the [MJ] network" to determine "the ancestral node" by "comparing the network nodes with suitable outgroups" (p. 29, user guide, fluxus-engineering.com), which has been implemented in some studies (e.g. Sakaguchi et al., 2012), this procedure merely links the "outgroup" sequence (i.e. non-conspecific) to the *most similar* haplotype of the already produced "ingroup" network. Hence, it neither roots the ingroup topology nor polarizes character transformations. The direct addition of an outgroup sequence into the network construction process (i.e. not through the rooting option) is likely to yield unresolved and extremely confusing networks with complex 3-D cycles and multiple median vectors (our unpublished analyses).

ian network (Bandelt et al., 1995), MJ can handle large datasets, as well as multistate data, such as amino acid sequences, very rapidly. In an attempt to create an intermediate-sized network, MJ combines the minimum spanning network and quasi-median network algorithms (the former produces too few nodes associated with a multiple alignment of sequences, and the latter produces too many; Huson et al., 2010). BEA99 (p. 37) developed the following reasoning:

> The MJ method begins with the minimum spanning trees, all combined within a single (reticulate) network. Aiming at parsimony, we subsequently add a few consensus sequences (i.e., median vectors, or Steiner points) of three mutually close sequences at a time. These median vectors can be biologically interpreted as possibly extant unsampled sequences or extinct ancestral sequences. The median operation, also referred to as "Steinerization" in mathematics (in which the most parsimonious realizations of MP trees are called Steiner trees; see Hwang et al., 1992), is basic to all fast MP heuristic algorithms, although it is typically applied in a very restricted ("greedy") manner in order to arrive at a single tree (Farris, 1970). In contrast, the unconstrained use of the median operation eventually generates the so-called full quasimedian network…, which normally harbors all optimal trees, as well as numerous suboptimal trees…With MJ, we take care that at each stage only those median vectors which have a good chance of appearing as branching nodes in an MP tree are generated by considering only triplets of sequences for which one sequence is linked to the other two in the network under processing. An additional ranking of these candidate triplets according to a distance score (as proposed by Tateno, 1990) allows further refinement of the triplet selection. After each round of median generation, the process restarts with the thus enlarged set of sequences.

Minimum spanning network. The MJ method starts with the generation of a minimum spanning network (BEA99). For a given set, S, of DNA sequences (s_1, s_2, s_3…, s_i), the algorithm requires a multiple sequence alignment, A, with infrequent ambiguous states and with no recombination. The data are the basis of a distance matrix, D, on S. To calculate the distance, d, between two sequences in A, the algorithm employs the Hamming distance, H, which is the number of differences between equal-length sequences [or to define D, then $H(s_x, s_y)$ in A]. Distance values between sequences are then increasingly ordered ($d_1 < d_2 < d_3 < … < d_i$). If character state differences between sequences (not "character changes" as claimed by BEA99, p. 39) are unequally weighted, then $d(s_x, s_y)$ in A to define D is given by the sum of weights, ω, of all different states between s_x and s_y (Fig. 4a), or

$$d(s_x, s_y) = \sum \omega_{(s_x, s_y)}^{\text{i-th}},$$

where i-th denotes any position at which s_x and s_y differ. This "weighted Hamming distance" of MJ (BEA99, p. 39) is better termed an "unequally
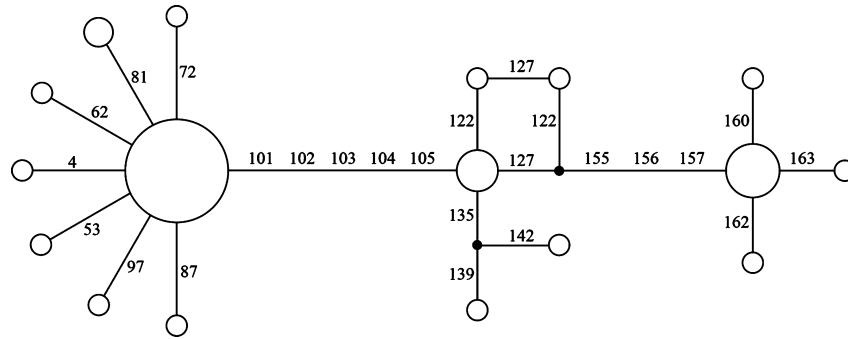
Fig. 3. A median-joining network. Note the presence of cycles due to the absence of direction (square). Each circle represents a unique haplotype where the diameter is proportional to the number of DNA sequences represented. Integers on each edge denote the position of nucleotides within the sequence that *differ* between haplotypes. Small solid circles indicate median vectors.
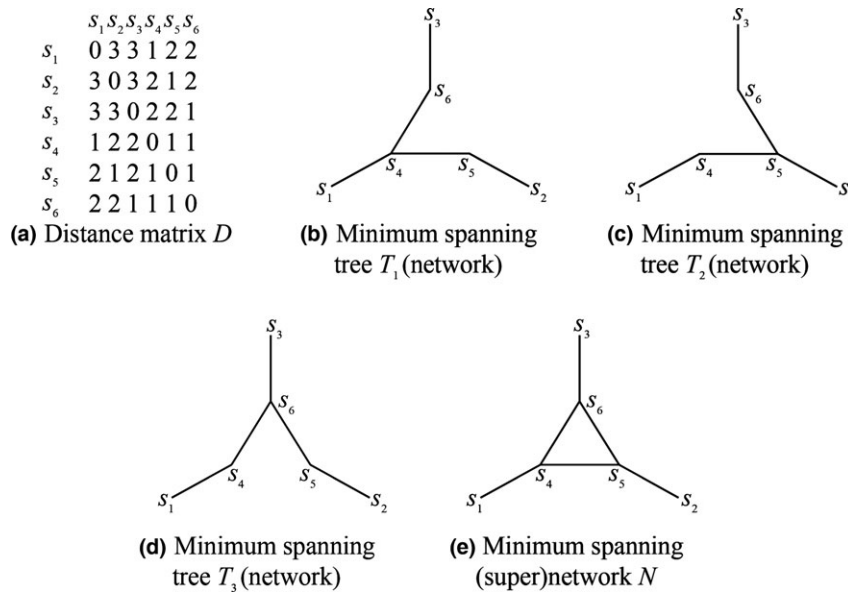

Fig. 4. The distance matrix D on $S = \{s_1, s_2, s_3 \ldots, s_i\}$ shown in (a) gives rise to three different minimum spanning trees (networks), shown in (b), (c) and (d), respectively. The corresponding minimum spanning (super)network N is shown in (e). Modified from Huson et al. (2010).

weighted Hamming distance". It would have been more precise because in the Hamming distance, differences are equally weighted. Ambiguous states are specified via a comparison "with the definite states of the other minimally distant sequences", and arbitrarily assigned by setting "the most common definite state of these sequences" (BEA99, p. 39).

As for the minimum spanning network itself, consider the graph $G = (V, E)$, where V represents the set of vertices (nodes), v, therefore $V = S$ in this case, and E is the set of edges (branches), e, containing all possible edges between any two nodes in V. In a *spanning tree* of G, all vertices are connected, and, as with trees in general, cycles cannot be formed. Note that the usage of "tree" in this context (and in mathematics) refers to a connected—not rooted—graph. If the edges have weights [$\omega(e)$] that reflect D, i.e. $\omega(e) = d(s_x, s_y)$

for every e in E, then $\omega(G)$ is given by the sum of the weights of all edges in G associated with A [$G = (V, E, \omega)$]; i.e. G is a "weighted graph" [Wheeler (2012), or a "distance graph" *sensu* Huson et al. (2010)]. In this case, the *minimum spanning tree* for G is the spanning tree (T) that connects all vertices of G and that minimizes the sum of $\omega(e)$ given by

$$\omega(T) = \sum e \text{ in } T^{\omega(e)}$$

(Huson et al., 2010; Fig. 4, e.g. b). In other words, it is the spanning tree whose weight is minimal (Wheeler, 2012).

Kruskal's (1956) and Prim's (1957) algorithms can construct minimum spanning trees. MJ analysis is based on the former. It can obtain different minimum spanning trees because it processes all edges that have

the same weight consecutively in some arbitrary order. This procedure serves as an implicit "tie-breaking" rule. Thus, different input orders can produce different, but equally optimal solutions (Fig. 4b–d; Kruskal, 1956; Huson et al., 2010, p. 229). The *minimum spanning network*, *N*, for *A* is the subgraph of the weighted graph *G* whose *E* is given by the union of the *E* of all minimum spanning trees generated (Huson et al., 2010; Fig. 4e). Simply stated, *N* is the union of all minimum spanning trees by dropping the tie-breaking rule of Kruskal's algorithm (BEA99). Cycles may be created initially because equally optimal solutions can be included and displayed simultaneously (Fig. 4e: triangle).

The tolerance parameter Δ (ε in BEA99) can be specified to restrict the distance values of accepted weighted edges into *N*. Naturally, parameter Δ in *N* is, by definition, = 0. If, instead, Δ is increased (i.e. >0), the distance criterion will be relaxed and thus *N* will also contain all of those heavier edges of *G* whose weights do not surpass the heaviest weight in *N* by more than Δ. The resulting graph is known as the *relaxed minimum spanning network* (N_Δ) (Huson et al., 2010). In practice, however, parameter Δ is usually set = 0. The obtained minimum spanning network is the starting point for the last process to build the final MJN, the addition of median vectors.

The MJ algorithm and the resulting MJN. Based on the quasi-median operation (Huson et al., 2010), but guided by the minimum spanning network to avoid the full quasi-median resolution, the MJ algorithm produces "additional…sequence types" (BEA99) from existing ones. Specifically, it generates consensus sequences. The addition of extra vertices and associated weighted edges to the minimum spanning trees may further reduce the overall weight of the graph (Wheeler, 2012), a process known as Steinerization. User-specified Δ affects the construction of the minimum spanning network, and it governs the generation of median vectors (i.e. the additional sequences). Increasing Δ widens the search for potential new median vectors. However, because this parameter is usually set to zero, only minimal cost connections (i.e. vertices connected by minimal weighted edges) are considered. Basically, triplets of sequence types, where there are at least two feasible edges among them, are used to generate a median vector (or Steiner point; Fig. 3: small solid circles). These median vectors are added to the original pool of sequence-types, and the minimum spanning network is then recalculated with the newly enlarged set of sequences. Iterations occur until no further median vectors can be generated. Consequently, cycles are formed. Either the original cycles may be modified or new ones may be produced (Fig. 3: square). The final

product (the MJN) shows all feasible links in minimal cost connections plus Δ at most (Fig. 3).

Discussion and implications. Although BEA99 devoted most of their attention to median generation, the construction of minimum spanning networks is central to the MJ method. Therefore, the construction of MJNs is based entirely on a measure of similarity of DNA sequences. The reliance on a distance-based technique was described operationally and mentioned recurrently throughout their paper, but BEA99 neither addressed this topic explicitly nor discussed its implications. This suggests that they favoured similarity-based, phenetic methods, which are unquestionably the most criticized approaches by (most) systematists. Alternatively, they may have failed to distinguish similarity from character transformation as the basis for delimiting groups. This oversight has led some phylogeneticists to mistakenly use phenetic approaches to infer phylogenetic relationships (Grant and Kluge, 2004). Subtle, but significant, confusion supports this possibility (BEA99, p. 39, italics added):

> The simplest way to obtain a distance measure between two sequences is to count the *number of character differences* (the "Hamming distance"). As a refinement, *we may also weight character changes*, albeit only in a symmetrical fashion…

In one way or another, all phylogenetic methods [maximum parsimony (MP), maximum likelihood and Bayesian inference] aim to minimize character transformations, but assumptions about character evolution employed in MJ analysis rely on similarity alone. Likewise, due to the spread of this method, MJ practitioners have been led to the assumption that a cluster in the final network is a group of closely related subjects, but these relationships are based on overall similarity of sequences, which is the operational basis of phenetics. Indeed, even if the MJ method employed outgroup rooting, it would behave like phenetic clustering. Unfortunately, this type of method is used frequently.

Median generation, as interpreted by BEA99, has its own problems. In an attempt to highlight the benefits of their method, they appealed to parsimony as a principle of MJ analysis [p. 37 (abstract)]:

> We present a method … for constructing networks … that combines features of Kruskal's algorithm for finding minimum spanning trees by favoring short connections, and Farris's maximum-parsimony (MP) heuristic algorithm, which sequentially adds new vertices called "median vectors", except that our MJ method does not resolve ties.

BEA99 continued (p. 37):

> Aiming at parsimony … [t]he median operation, … (in which the most parsimonious realizations of MP trees are called Steiner trees), … is basic to all fast MP heuristic algorithms, although it is typically applied in a very restricted ("greedy")

manner in order to arrive at a single tree (Farris, 1970)...
With MJ, we take care that at each stage only those median
vectors which have a good chance of appearing as branching
nodes in an MP tree are generated by considering only triplets
of sequences for which one sequence is linked to the other
two in the network under processing.

Likewise, they justified the addition of median vectors
as follows (p. 38):

> [T]he minimum spanning network is of little direct use for
> representing genetic data, since in general a minimum span-
> ning tree is far from being most parsimonious. It serves, how-
> ever, as a good point of departure in each recursive step of
> our MJ network construction for generating additional
> inferred sequence types which reduce tree length.

The interpretation of MP by BEA99 differed from
that of Farris (1970). Invoking principles of parsimony
does not validate a phenetic technique as being a phy-
logenetic method, and the best Steiner trees are not
the most parsimonious trees. Again, "tree" in the MJ
context refers to a connected—not rooted—graph, or
network. The generation of median vectors and the
subsequent addition of extra vertices to the minimum
spanning tree(s) may reduce the overall weight of the
graph(s), which is then referred to as a Steiner tree(s).
Consequently, the overall weight of the minimum
spanning networks and MJNs may also be reduced.
This Steinerization process in MJ analysis serves to
obtain MP, and the resulting best (unrooted) Steiner
tree is what BEA99 assumed to be the most parsimo-
nious tree. Although reducing the weight of graphs
could constitute a form of Occam's Razor, this form
of MP does not empower MJ.

A Prim Network is an unrooted, connected graph in
which the set of nodes (i.e. vertices) is identical to that
of observed taxa (operational taxonomic units, OTUs),
in this case DNA sequences. Thus, no median nodes
(vectors) or hypothetical taxonomic units (HTUs) are
constructed (Farris, 1970). This immediately leads to
what is usually referred to as a Steiner-type problem
in systematics: extant taxa cannot be ancestors of
other extant taxa (Wheeler, 2012). Wagner or Steiner
networks (also undirected, connected graphs), in turn,
allow for the addition of extra vertices (i.e. HTUs or
Steiner points) and associated edges. However, the
Steiner problem expands on them because OTUs and
HTUs are placed indiscriminately on the network in
some order that is determined by a given cost function
and the absence of direction. In sharp contrast, a
(rooted) Steiner tree is a minimum cost tree with a set
of terminal and internal vertices (or Steiner points),
and is therefore a Wagner tree (Farris, 1970). In Wag-
ner trees, OTUs are confined to terminal nodes (tips
or leaves) and HTUs are placed at inner nodes. Thus,
Wagner or Steiner trees overcome the Steiner problem
(Platnick, 1977). Under this scenario, all "trees" in MJ
are actually Prim or Steiner networks. Thus, minimum

spanning networks are best defined as "minimum
spanning super-networks" (Fig. 4e). Although
BEA99's interpretation of MP goes beyond the theo-
retical underpinnings of network- and tree-building
differences and implications, they demark a relevant
starting point for understanding the essential distinc-
tion between median construction of the MJ algorithm
and that of Farris (1970): character optimization.

The Wagner method (Farris, 1970; additive charac-
ter optimization) is a modified version of the original
Wagner procedure (Wagner, 1961). It builds a single
branching diagram by adding OTUs sequentially (one
at a time) in an arbitrary order. However, the optimal
position on a growing tree is determined by choosing
the immediately best option. Specifically applied for
additive characters, the algorithm finds one optimiza-
tion, but not all, as pointed out later by Farris (1983)
and Goloboff (1993). More parsimonious solutions are
found by optimizing novel character combinations
independently to inner nodes, creating HTUs. The
resulting tree is commonly used as a starting point for
several other heuristics (e.g. random addition
sequences, RAS) to find more parsimonious solutions
(Wheeler et al., 2006; Wheeler, 2012). Different
approaches have also been proposed to find all possi-
ble optimizations of characters (e.g. Goloboff, 1993),
but their description is beyond the scope of this paper.
The simple Wagner algorithm (Farris, 1970), which
BEA99 made reference to when claiming MP for MJ
analysis, generates medians (HTUs) based on the char-
acter states of the OTUs and other HTUs. Initially,
the addition and placement of OTUs is determined by
the advancement index, which establishes a rank order,
and the interval distance formula, a relation on the
character-states between HTUs and OTUs. New
HTUs to connect OTUs to the branching diagram are
formed through the median-state property, which spec-
ifies optimal HTUs one character-state at a time (Far-
ris, 1970). The Wagner method uses patristic distances
(character-state transformations, number of steps, tree-
length) of a phylogenetic hypothesis (cladogram) to
explain the observed character variation (Kluge and
Grant, 2006; also see Farris, 1967), and not similarity
or phenetic differences. Whereas the Wagner algorithm
aims to minimize character transformations (i.e. a
character-based method), the MJ algorithm is gov-
erned completely by similarity. Thus, Farris' algorithm
and the MJ method are similar only in their sequential
addition of new vertices to a diagram under construc-
tion (i.e. a graphic procedure from opposite
approaches).

BEA99 stressed that median vectors can be inter-
preted biologically as existing unsampled or extinct
ancestral sequences (i.e. they can represent missing
intermediates; Fig. 3). However, a median vector in an

MJ analysis is a majority-consensus-generated sequence and a mathematically drawn point in the final MJN that connects a triplet of sequences. The resulting "evolutionary paths in the form of cycles" (BEA99, p. 37) merely illustrates the failure of the algorithm to choose between alternative, equally optimal connections due to the modification of Kruskal's algorithm. Consequently, a cycle represents an analytical artefact rather than an evolutionary scenario (Salzburger et al., 2011).

Finally, BEA99 introduced a new meaning for homoplasy (p. 37): "[t]he MJ method…can be adjusted to the level of homoplasy by setting a [tolerance] parameter ε [Δ]", because (p. 39),

> In practice, the quasimedian network generated by the given data may be somewhat large due to homoplasy, such that only a portion should be heuristically constructed by carefully selecting triplets of sequences for median generation.

By homoplasy, BEA99 meant the multitude of nodes generated by the full quasimedian resolution, which can be constrained by implementing the MJ operation. The MJ analysis, like any other phenetic (not to mention undirected) method, does not recognize either homoplasy or homology, and it does not distinguish between them.

MJNs in the literature

To evaluate the impact of the MJ analysis on evolutionary biology, we conducted a meta-analysis of studies that employed the MJ method to infer evolutionary relationships among organisms. Over the last 15 years, more than 3000 papers have cited BEA99 [ISI Web of Science; absolute number of citations (December, 2014)]. The number of citations increased exponentially from 118 (<5%) between 1999 and 2003 to over 2023 (>55%) between 2010 and 2014 (Fig. 1). We further examined an initial subset of 376 randomly chosen articles between 2006 (first year with more than 100 citations) and 2014 and without consideration of taxonomic group. Theoretical papers that did not employ the method were filtered out. Recent applications of MJNs included population structure analysis (e.g. Escobar-Gutiérrez et al., 2013), visualization of haplotype diversity (e.g. McCracken et al., 2013) and relationships (e.g. Castelin et al., 2012), complementary information of phylogenetic analyses (i.e. when phylogenetic trees could not provide "desired resolution"; e.g. Klütsch et al., 2012) and phylogenetic inference. We further evaluated the last-named application, although most discoveries also apply equally to the former three usages.

Among the 376 papers, 161 studies (42.8%) used MJNs primarily to infer "phylogenetic" (e.g. Zhu et al., 2013), "evolutionary" (e.g. Turchetto et al., 2014) or "genealogical" (e.g. Amaral et al., 2014) relationships, as stated in the Materials and Methods section of the articles. Although MJNs were mostly used to infer phylogenies based on intraspecific haplotypes (e.g. Fan et al., 2013), the MJ method was also used to infer interspecific relationships (e.g. Zou et al., 2013). Additional inappropriate use of terms was common in the literature.

Authors often misinterpreted similar haplotypes grouped through MJN as "clades" (e.g. Cunha et al., 2012), and the MJ product as "cladograms" (e.g. Cao et al., 2013). Clades are monophyletic groups, i.e. an ancestor and all its descendants, which is represented on a rooted cladogram by all terminals arising from a single node (Wheeler et al., 2006). A grouping produced by MJ is simply a crowd of haplotypes that are similar to each other (Fig. 3). Hence, it is more appropriate to refer to them as "clusters" or "haplogroups". Similarly, MJNs should be termed "unrooted networks" or "phenograms" only.

Confusion between MJNs and MP trees exist. For example, Malyarchuk et al. (2014) reported that optimal, shortest phylogenetic trees were reconstructed using the MP calculation in optional post-processing implemented in NETWORK (Polzin and Daneshmand, 2003). Unfortunately, this is untrue. MP calculations of NETWORK do not produce MP trees, but rather they identify median vectors and edges produced in a full MJN that are not contained in the shortest graphs, and switches them off in the final display (Fig. 3). Sometimes authors included MJ analyses along with traditional phylogenetic tree-reconstruction methods. For example, Bataille et al. (2013, p. 4199) inferred phylogenetic relationships using "maximum parsimony, maximum-likelihood, Bayesian inference, and Median-Joining Network methods".

One of the prerequisites for MJ analysis is recombination-free input data (BEA99). MJNs are not intended to detect recombination events and, indeed, such would be impossible. Notwithstanding, MJ was used to "infer recombination amongst haplotypes" (Arnott et al., 2013, p. 4).

Comparative studies including MJNs are over-interpreted in many cases to justify the use of the method. Cassens et al. (2003, 2005) and Wooley et al. (2008) examined the relative performance of MJNs compared with several other approaches. These have served as reference studies. Cassens et al. (2003) stated that the MJ method yielded the "best genealogy" because it required the least number of mutations to explain the data when compared with other unrooted networks. Furthermore, Cassens et al. (2005) suggested the MJ approach worked well when haplotypes were relatively distant, yet it occasionally failed to reconstruct

the "correct topology". In contrast, Wooley et al. (2008) reported that MJNs performed well when the substitution rate was low, but performed significantly less accurately when the substitution rate was high. Some authors exaggerated the findings from Cassens et al. (2003) by stating that MJNs have been shown to yield "the best-resolved genealogies relative to other rooting and network procedures" (Lin et al., 2012, p. e36334; Zhou et al., 2011, p. 331). Paupério et al. (2012), p. 6019) also justified use of MJNs by referring to Cassens et al. (2005) and Wooley et al. (2008) and stating "in the analysis of closely related sequences, *networks* are useful tools as they *can provide more information than a strict consensus tree* and still *present a reliable estimate of the true genealogy*" (emphasis added).

Conclusions

Other than fast computation and very attractive graphics, MJNs harbour no virtue for phylogenetic inference. MJNs are distance-based, unrooted branching diagrams with cycles that say nothing about the evolutionary history due to the absence of direction. MJ was introduced in 1999 and, in contrast to most scientific ideas, its application has spread rapidly through copying the methods of others, and, unfortunately, with little further scrutiny. We hope that the theoretical arguments presented here can reverse this trend.

Acknowledgements

References

Amaral, A.R., Lovewell, G., Coelho, M.M., Amato, G., Rosenbaum, H.C., 2014. Hybrid speciation in a marine mammal: the Clymene dolphin (*Stenella clymene*). PLoS ONE 9, e83645.

Arnott, A., Mueller, I., Ramsland, P.A., Siba, P.M., Reeder, J.C., Barry, A.E., 2013. Global population structure of the genes encoding the malaria vaccine candidate, *Plasmodium vivax* apical membrane Antigen 1 (*Pv*AMA1). PLoS Negl. Trop. Dis. 7, e2506.

Bandelt, H.J., Forster, P., Sykes, B.C., Richards, M.B., 1995. Mitochondrial portraits of human populations. Genetics 141, 743–753.

Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Bataille, A., Fong, J.J., Cha, M., Wogan, G.O., Baek, H.J., Lee, H., Min, M.S., Waldman, B., 2013. Genetic evidence for a high diversity and wide distribution of endemic strains of the pathogenic chytrid fungus *Batrachochytrium dendrobatidis* in wild Asian amphibians. Mol. Ecol. 22, 4196–4209.

Cao, L., Liang, X.-F., Du, Y., Zheng, H., Yang, M., Huang, W., 2013. Genetic population structure in *Siniperca scherzeri* (Perciformes: Siniperca) in China inferred from mitochondrial DNA sequences and microsatellite loci. Biochem. Syst. Ecol. 51, 160–170.

Cassens, I., Van Waerebeek, K., Best, P.B., Crespo, E.A., Reyes, J., Milinkovitch, M.C., 2003. The phylogeography of dusky dolphins (*Lagenorhynchus obscrus*): a critical examination of network methods and rooting procedures. Mol. Ecol. 12, 1781–1792.

Cassens, I., Mardulyn, P., Milinkovich, M.C., 2005. Evaluating intraspecific "network" construction methods using simulated sequence data: do existing algorithms outperform the Global Maximum Parsimony approach? Syst. Biol. 54, 363–372.

Castelin, M., Lorion, J., Brisset, J., Cruaud, C., Maestrati, P., Utge, J., Samadi, S., 2012. Speciation patterns in gastropods with long-lived larvae from deep-sea seamounts. Mol. Ecol. 21, 4828–4853.

Chung, F.R.K., 1986. Diameters of communication networks. Proc. Symp. Appl. Math. 34, 1–18.

Cunha, R.L., Coscia, I., Madeira, C., Mariani, S., Stefanni, S., Castilho, R., 2012. Ancient divergence in the trans-oceanic deep-sea shark *Centroscymnus crepidater*. PLoS ONE 7, e49196.

Delwiche, C.F., Palmer, J.D., 1996. Rampant horizontal transfer and duplication of rubisco genes in eubacteria and plastids. Mol. Biol. Evol. 13, 873–882.

Eldredge, N., Cracraft, J., 1980. Phylogenetic Patterns and the Evolutionary Process. Columbia University Press, New York.

Escobar-Gutiérrez, A., Soudeyns, H., Larouche, A., Carpio-Pedroza, J.C., Martinez-Guarneros, A., Vazquez-Chacon, C.A., Fonseca-Coronado, S., Yamasaki, L.H., Ruiz-Tovar, K., Cruz-Rivera, M., 2013. Vertical transmission of hepatitis C virus: a tale of multiple outcomes. Infect. Genet. Evol. 20, 465–470.

Fan, D.M., Yue, J.P., Nie, Z.L., Li, Z.M., Comes, H.P., Sun, H., 2013. Phylogeography of *Sophora davidii* (Leguminosae) across the 'Tanaka-Kaiyong line', an important phytogeographic boundary in Southwest China. Mol. Ecol. 22, 4270–4288.

Farris, J.S., 1967. The meaning of relationship and taxonomic procedure. Syst. Zool. 16, 44–51.

Farris, J.S., 1970. Methods for computing Wagner trees. Syst. Zool. 19, 83–92.

Farris, J.S., 1972. Estimating phylogenetic trees from distance matrices. Am. Nat. 106, 645–668.

Farris, J.S., 1982. Outgroups and parsimony. Syst. Zool. 31, 328–334.

Farris, J.S., 1983. The logical basis of phylogenetic analysis. In: Platnick, N.I., Funk, V.A. (Eds.), Advances in Cladistics. Columbia University Press, New York, pp. 277–302.

Farris, J.S., Kluge, A.G., Mickevich, M.F., 1982. Phylogenetic analysis, the monothetic group method, and myobatrachid frogs. Syst. Zool. 31, 317–327.

Gangloff, B., Zino, F., Shirihai, H., González-Solís, J., Couloux, A., Pasquet, E., Bretagnolle, V., 2013. The evolution of North-East Atlantic gadfly petrels using statistical phylogeography. Mol. Ecol. 22, 495–507.

Goloboff, P.A., 1993. Character optimization and calculation of tree lengths. Cladistics 9, 433–436.

Grant, T., Kluge, A.G., 2004. Transformation series as an ideographic character concept. Cladistics 20, 23–31.

Grant, T., Faivovich, J., Pol, D., 2003. The perils of 'point-and-click' systematics. Cladistics 19, 275–285.

Griffiths, R.C., Marjoram, P., 1996. Ancestral inference from samples of DNA sequences with recombination. J. Comput. Biol. 3, 479–502.

Huson, D.H., Rupp, R., Scornavacca, C., 2010. Phylogenetic Networks: Concepts, Algorithms and Applications. Cambridge University Press, Cambridge.

Hwang, F.K., Richards, D.S., Winter, P., 1992. The Steiner Tree Problem. North-Holland, Amsterdam.

Kluge, A.G., Grant, T., 2006. From conviction to anti-superfluity: old and new justifications of parsimony in phylogenetic inference. Cladistics 22, 276–288.

Klütsch, C.F.C., Manseau, M., Wilson, P.J., 2012. Phylogeographic analysis of mtDNA data indicates postglacial expansion from multiple glacial refugia in Woodland caribou (*Rangifer tarandus caribou*). PLoS ONE 7, e52661.

Kruskal, J.B., 1956. On the shortest spanning subtree of a graph and the traveling salesman problem. Proc. Am. Math. Soc. 7, 48–50.

Lin, L.-H., Qu, Y.-F., Li, H., Zhou, K.-Y., Ji, X., 2012. Genetic structure and demographic history should inform conservation: Chinese cobras currently treated as homogenous show population divergence. PLoS ONE 7, e36334.

Linder, C.R., Rieseberg, L.H., 2004. Reconstructing patterns of reticulate evolution in plants. Am. J. Bot. 91, 1700–1708.

Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536.

Malyarchuk, B., Derenko, M., Mikhailova, E., Denisova, G., 2014. Phylogenetic relationships among *Neoechinorhynchus* species (Acanthocephala: Neoechinorhynchidae) from North-East Asia based on molecular data. Parasitol. Int. 63, 100–107.

McCracken, K.G., Wilson, R.E., Peters, J.L., Winker, K., Martin, A.R., 2013. Late Pleistocene colonization of South Georgia by yellow-billed pintails pre-dates the Last Glacial Maximum. J. Biogeogr. 40, 2348–2360.

Moret, B.M.E., Nakhleh, L., Warnow, T., Linder, C.R., Tholse, A., Padolina, A., Sun, J., Timme, R., 2004. Phylogenetic networks: modeling, reconstructibility, and accuracy. IEEE/ACM Trans. Comput. Biol. Bioinform. 1, 13–23.

Paupério, J., Herman, J.S., Melo-Ferreira, J., Jaarola, M., Alves, P.C., Searle, J.B., 2012. Cryptic speciation in the field vole: a multilocus approach confirms three highly divergent lineages in Eurasia. Mol. Ecol. 21, 6015–6032.

Planet, P.J., Kachlany, S.C., Fine, D.H., DeSalle, R., Figurski, D.H., 2003. The widespread colonization island of *Actinobacillus actinomycetemcomitans*. Nat. Genet. 34, 193–198.

Platnick, N.I., 1977. Cladograms, phylogenetic trees, and hypothesis testing. Syst. Zool. 26, 438–442.

Polzin, T., Daneshmand, S.V., 2003. On Steiner trees and minimum spanning trees in hypergraphs. Oper. Res. Lett. 31, 12–20.

Prim, R.C., 1957. Shortest connection networks and some generalizations. Bell Syst Tech. J. 36, 1389–1401.

Sakaguchi, S., Qiu, Y.X., Liu, Y.H., Qi, X.S., Kim, S.H., Han, J., Takeuchi, Y., Worth, J.R., Yamasaki, M., Sakurai, S., Isagi, Y., 2012. Climate oscillation during the Quaternary associated with landscape heterogeneity promoted allopatric lineage divergence of a temperate tree *Kalopanax septemlobus* (Araliaceae) in East Asia. Mol. Ecol. 21, 3823–3838.

Salzburger, W., Ewing, G.B., von Haeseler, A., 2011. The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. Mol. Ecol. 20, 1952–1963.

Song, Y.S., Hein, J., 2005. Constructing minimal ancestral recombination graphs. J. Comput. Biol. 12, 147–169.

Tateno, Y., 1990. A method for molecular phylogeny construction by direct use of nucleotide sequence data. J. Mol. Evol. 30, 85–93.

Templeton, A.R., Crandall, K.A., Sing, C.F., 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132, 619–633.

Turchetto, C., Fagundes, N.J., Segatto, A.L., Kuhlemeier, C., Solís Neffa, V.G., Speranza, P.R., Bonatto, S.L., Freitas, L.B., 2014. Diversification in the South American pampas: the genetic and morphological variation of the widespread *Petunia axillaris* complex (Solanaceae). Mol. Ecol. 23, 374–389.

Wagner, W.H., 1961. Problems in the classification of ferns. In: Recent Advances in Botany. University of Toronto Press, Toronto, pp. 841–844.

Wheeler, W.C., 2012. Systematics: A Course of Lectures. Wiley-Blackwell, Oxford.

Wheeler, W.C., 2014. Phyletic groups on networks. Cladistics 30, 447–451.

Wheeler, W., Aagesen, L., Arango, C.P., Faivovich, J., Grant, T., D'Haese, C., Janies, D., Smith, W.L., Varón, A., Giribet, G., 2006. Dynamic Homology and Phylogenetic Systematics: A Unified Approach Using POY. American Museum of Natural History, New York.

Wooley, S.M., Posada, D., Crandall, K.A., 2008. A comparison of phylogenetic network methods using computer simulation. PLoS ONE 3, e1913.

Zhou, C., Li, M., Yuan, K., Hu, N., Peng, W., 2011. Phylogeography of *Ascaris lumbricoides* and *A. suum* from China. Parasitol. Res. 109, 329–338.

Zhu, L., Hu, Y., Qi, D., Wu, H., Zhan, X., Zhang, Z., Bruford, M.W., Wang, J., Yang, X., Gu, X., Zhang, L., Zhang, B., Zhang, S., Wei, F., 2013. Genetic consequences of historical anthropogenic and ecological events on giant pandas. Ecology 94, 2346–2357.

Zou, J., Sun, Y., Li, L., Wang, G., Yue, W., Lu, Z., Wang, Q., Liu, J., 2013. Population genetic evidence for speciation pattern and gene flow between *Picea wilsonii*, *P. morrisonicola* and *P. neoveitchii*. Ann. Bot. 112, 1829–1844.