





## Opinion

Next-generation cophylogeny: unravelling  
eco-evolutionary processes

Isabel Blasco-Costa <sup>1,2,6,\*,@</sup> Alexander Hayward <sup>3,7,@</sup>  
Robert Poulin <sup>4,8</sup> and Juan A. Balbuena <sup>5,9,@</sup>

A fundamental question in evolutionary biology is how microevolutionary processes translate into species diversification. Cophylogeny provides an appropriate framework to address this for symbiotic associations, but historically has been primarily limited to unveiling patterns. We argue that it is essential to integrate advances from ecology and evolutionary biology into cophylogeny, to gain greater mechanistic insights and transform cophylogeny into a platform to advance understanding of interspecific interactions and diversification more widely. We discuss key directions, such as incorporating trait reconstruction and considering multiple scales of network organization, and highlight recent developments for implementation. A new quantitative framework is proposed to allow integration of relevant information, such as quantitative traits and assessment of the contribution of individual mechanisms to cophylogenetic patterns.

## Symbiotic associations in the light of coevolution

Biotic interactions pervade all biological systems and, since no species evolves in isolation, it can be claimed that nothing in evolution makes sense except in the light of **coevolution** (see [Glossary](#)) [1]. This is especially evident in **symbiosis** (**parasitism**, **mutualism**, and **commensalism**). All animals and plants, and many other organisms, host symbionts that contribute to their genetic variation with genes and associated functions, and/or impact their fitness [2,3]. Symbiosis plays key roles in ecosystem functioning and stability and is a major driver of evolutionary novelty and ecological diversity [4]. Symbiotic interactions often involve reciprocal selection dynamics and other microevolutionary processes [5–7]. Despite recognition of these important influences, and the central role of ecology and evolution in the organization and maintenance of biodiversity [8], understanding of how symbiotic interactions bring about macroevolutionary change remains incomplete.

Illuminating how microevolutionary processes lead to macroevolutionary change requires being able to disentangle how ecological and evolutionary processes influence lineage diversification [8]. **Cophylogeny** provides an appropriate framework to address this fundamental question in symbiotic associations. Classical cophylogeny evaluates the dependency of evolutionary histories of two groups of organisms based on extant (ecological scale) relationships to discover constraints and predictability in their relationships and influences on their diversification [9] ([Box 1](#)). More recently, interest in the phylogenetic patterns behind species associations has leaked from evolutionary ecology into general ecology as a way to understand species assembly in communities of associated organisms [10–12]. Over the last 40 years, studies have reported a variety of cophylogenetic patterns from different symbiotic relationships and much progress has been made in the development of analytical tools (e.g., [13–15]) (Figure S1 in the supplemental information online). However, linking patterns to mechanisms remains the major outstanding challenge of cophylogeny [16–18].

## Highlights

Cophylogeny provides an appropriate setting to untie how the ecological and evolutionary facets of species interactions operate.

The study of phylogenetic agreement between the evolutionary histories of two groups of symbionts started as a means to analyze constraints and predictability in their interactions and codiversification.

More recently, it has attracted the attention of other areas in a quest to understand phylogenetic signal in community assembly and geographical distributions of species.

Whilst the field has fostered the development of a battery of tools to elucidate cophylogenetic patterns in symbiotic associations, linking patterns to mechanisms remains the major outstanding challenge.

Incorporating current trends in eco-evolutionary research could propel forward cophylogenetic principles and analyses.

<sup>1</sup>Department of Invertebrates, Natural History Museum of Geneva, PO Box 6434, CH-1211 Geneva 6, Switzerland

<sup>2</sup>Department of Arctic and Marine Biology, UiT The Arctic University of Norway, Langnes, PO Box 6050, 9037 Tromsø, Norway

<sup>3</sup>Centre for Ecology and Conservation, University of Exeter, Penryn Campus, Penryn, Cornwall, Exeter, TR10 9FE, UK

<sup>4</sup>Department of Zoology, University of Otago, PO Box 56, Dunedin, New Zealand

<sup>5</sup>Cavanilles Institute of Biodiversity and Evolutionary Biology, University of Valencia, PO Box 22085, 46071 Valencia, Spain

### Box 1. Overview of cophylogenetic methodology

Cophylogenetic methods fall into two categories: **events-based methods** and **global fit methods** [14, 15, 62]. The former map symbiont phylogeny onto host phylogeny to estimate coevolutionary events. The two phylogenies are considered more congruent than expected by chance, if the estimated number of coevolutionary events exceeds a significance threshold when considering the distribution of events obtained by randomization (the *P* value equals the proportion of times that randomized symbiont trees map to the host tree at least as well as the original symbiont tree). In some implementations (e.g., [63]), constraints or costs may be set during the mapping stage to model different biological scenarios. Global fit methods evaluate the congruence between host and symbiont trees by considering genetic distances representing functions of host and symbiont phylogeny. The extent of congruence between the two trees is assessed by permutating the dataset many times to determine the significance of the observed result [64, 65].

A drawback of events-based methods is that even with a small number of taxa the number of equally parsimonious solutions can be exceedingly high and they are strongly dependent on a good estimation of the set of costs considered. Additionally, estimation of congruence between the observed trees is typically based on maximizing cospeciation events, which is not always biologically realistic [15, 39, 66]. Whilst there are well-documented solutions to these problems [67], they are rarely implemented in practice. Global fit methods are typically more computationally efficient and allow the evaluation of phylogenetic uncertainty and the contribution of individual host–symbiont links to overall congruence. However, they lack the explanatory power of their events-based counterparts [64].

Both approaches are only as good as the information supplied to them. Consequently, one must make sure that all relevant taxa have been sampled and that species have been delimited accurately with a proper account of cryptic diversity (a common issue in symbiont taxa; see [68]). For instance, phylogenetic congruence may be overestimated when closely related host species are associated with generalist (cryptic) symbionts [39]. New approaches that identify missing links in host–symbiont interaction networks offer a means by which to evaluate the effect of incomplete sampling, particularly when considering less well studied taxa [69, 70].

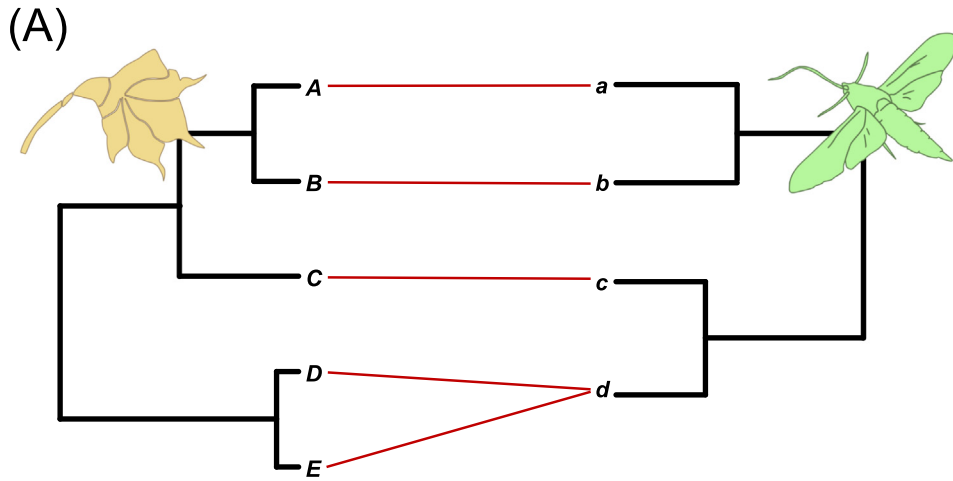
<sup>6</sup><https://blasco costa.ch>  
<sup>7</sup>[www.jumpinggenes.org](http://www.jumpinggenes.org)  
<sup>8</sup>[www.otago.ac.nz/parasitegroup/home.html](http://www.otago.ac.nz/parasitegroup/home.html)  
<sup>9</sup>[www.uv.es/~balbuena/research/research.html](http://www.uv.es/~balbuena/research/research.html)

\*Correspondence:  
[isabel.blasco-costa@ville-ge.ch](mailto:isabel.blasco-costa@ville-ge.ch)  
 (I. Blasco-Costa).  
 @Twitter: @PeerLab\_NHMG  
 (I. Blasco-Costa), @zanderhayward  
 (A. Hayward), and @JA\_Balbuena  
 (J. A. Balbuena).

One likely reason for this situation is the lack of alignment of cophylogeny with recent advances in eco-evolutionary theory and approaches. Consequently, our aim is to illustrate how the incorporation of current trends can fruitfully contribute to cophylogenetic principles and analyses. These include study of the evolution of phenotypic traits by means of phylogenetic comparative methods [19], assessing ecological interactions based on network theory [20], taking advantage of genomic data, consideration of intraspecific heterogeneity, and progress in dated phylogenies. Embracing this change will promote more mechanistic interpretations of cophylogenetic patterns and transform cophylogeny into a powerful platform to advance our understanding of the ecology and evolution of symbiosis.

### Integration of traits in cophylogeny

Classical cophylogenetic analysis compares the phylogenies of interacting partners based on observed extant interactions between them. The latter are often presented as a binary **host–symbiont interaction matrix**, where 0s and 1s codify, respectively, the absence or presence of interactions between taxa in nature (Figure 1A,B). While this approach provides a basis for investigating cophylogenetic relationships, it ignores variation in the strength of associations among interacting partners and thus precludes a deeper assessment of cophylogenetic processes. Many symbionts are not strict specialists, but generalists that utilize a range of host taxa, and interactions are often more frequent or intense for certain subsets of hosts (i.e., principal hosts) [21, 22]. Thus, to better reflect host–symbiont relationships and convey their strength, the interaction matrix should be quantitative. **Interaction traits** that measure a property of the two-sided interaction (host–symbiont) should be selected for this purpose (see later for one-sided interaction traits). For instance, in plant–pollinator relationships, the frequency of visits represents both the ability of the symbiont to locate the host and that of the host to attract the symbiont [23]. Likewise, in host–parasite systems, measures of prevalence, abundance, or virulence that result from the interaction between parasite exploitation strategy and host tolerance/resistance, also qualify as two-sided interaction traits [24]. A quantitative host–symbiont interaction matrix (Figure 1C), or several matrices (e.g., [25]), each with a different



(B)

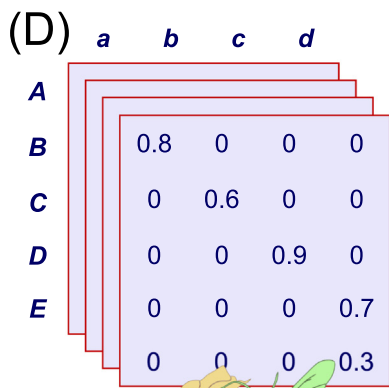
	a	b	c	d
A	1	0	0	0
B	0	1	0	0
C	0	0	1	0
D	0	0	0	1
E	0	0	0	1

Host-symbiont association matrix (binary)

(C)

	a	b	c	d
A	0.8	0	0	0
B	0	0.6	0	0
C	0	0	0.9	0
D	0	0	0	0.7
E	0	0	0	0.3

Host-symbiont association matrix (quantitative)



Multiple host-symbiont association matrices

(E)

	a	b	c	d
A	0.9	0.6	0.2	0.4
B	0.3	0.7	0	0
C	0.6	0	0.9	0
D	0.7	0	0	0.8
E	0.4	0	0	0.5

Host-symbiont codependence matrix

Glossary

**Bipartite network:** a network whose vertices (or nodes) are divided into two types, A and B, in which every edge connects a vertex of type A to one or more of type B and vice versa. No edges connect vertices within each type.

**Coadaptive codiversification:** reciprocal diversification of interacting lineages caused by coadaptation.

**Coevolution:** process of reciprocal evolutionary change caused by the interactions among species.

**Commensalism:** symbiotic relationship in which one of two partner species benefits and the other shows no apparent beneficial or harmful effect.

**Cophylogenetic signal:** higher congruence between the phylogenies of symbiotic partners than expected by chance.

**Cophylogeny:** the study of concordance between the phylogenies and the interactions of two groups of associated species.

**Cophylospace:** a cophylogenetic framework linking cophylogenetic patterns with processes via their representation in a three-dimensional space defined by the cophylogenetic signal and the associations between the bipartite network of symbiotic partners and their respective phylogenies.

**Cospeciation:** process of concomitant speciation in two or more interacting taxa. Can also be referred to as 'codivergence'.

**Diffuse coevolution:** evolution driven by interactions among several species in a community.

**Events-based methods:** cophylogenetic methods that attempt to find the best supported reconstructions of host and symbiont phylogenies using a set of events including cospeciation, duplication, host switching, and loss.

**Global fit methods:** cophylogenetic methods that test the overall congruence between host and symbiont phylogenies.

**Host switching:** (a.k.a. host shift); colonization of a novel host species by a symbiont.

**Host-symbiont interaction matrix:** structure of the associations between two groups of interacting partners.

**Interaction trait:** a trait that captures a property of the two-sided interaction between a host and a symbiont.

**Modularity:** a measure of network structure that quantifies the strength of the organization of the network into

interaction trait (Figure 1D), could be used as an input for cophylogenetic methods. Interest in such traits has been noted in studies evaluating **phylogenetic signal** in ecological communities [10,21], but not in classical cophylogenetic studies. The use of one or several quantitative traits will require adjustments that are within the reach of current cophylogenetic tools.

The importance of evolutionary history and ecological interactions in understanding the diversity of clades, communities, and their coevolution, has fostered interest in phylogenetic trait reconstruction as a link between micro- and macroevolution (e.g., [8,26]). We advocate for the inclusion of putative coevolving traits within cophylogenetic settings as a means to identify key features of the interacting species. In addition to two-sided interaction traits, additional traits, present in one or other symbiotic partner, are of interest. For instance, the length of mouth parts in pollinators reflects floral characteristics (e.g., [27,28]). Similarly, the morphology of parasite attachment organs often correlates with features of the host substrate [29,30]. In such cases, the morphological match between the partners may itself be considered an interaction trait. Contrary to two-sided interaction traits, one-sided traits evolve along the branches of one symbiotic partner only, although their evolution may depend on phylogeny and traits associated with the other partner. A codependence matrix accounting for the relationship between the traits on the symbiotic partners could be used as the interaction matrix (Figure 1E). Manceau *et al.* [31] developed an elegant, though as yet rarely applied, approach to consider one-sided interaction traits that evolve in a codependent manner through symbiont trees. The analysis of one-sided interaction traits holds promise to pinpoint the characters contributing to species specialization or promoting spill-over to new hosts and offers to facilitate interpretations of **cospeciation** and other processes inferred through cophylogenetic approaches.

### Focus on scale and network organization

Combining phylogenetic information on symbiotic partners and their structured network of interactions can be used to test alternative diversification scenarios and other evolutionary hypotheses [32]. Explicit consideration of the scale(s) at which these analyses are carried out could benefit future studies, because examining **cophylogenetic signal** across scales would offer the advantage of revealing the evolutionary processes operating at each level of organization [11].

Cophylogenetic studies are often carried out globally (i.e., they use phylogenetic and relational data collected over large spatial scales) (e.g., [33,34]). More rarely, authors have assessed potential differences in cophylogenetic patterns between regional and global scales [35]. Since **phylogenetic congruence** can arise from matching of deep nodes, testing cophylogenetic signal across scales would serve to establish the geographical scale and timing at which coevolutionary processes have operated. This approach was adopted by Hutchinson *et al.* [11], who suggested that the community is the most relevant scale at which cophylogenetic signal characterizes interactions between species assemblages. Moreover, the integration of interaction traits proposed earlier probably makes most sense at lower (community/regional) levels.

Cophylogenetic signal also permeates to finer scales. Interactions between symbiotic partners can be described with a **bipartite network** that typically reveals nonrandom associations

modules of taxa (i.e., taxa that interact more frequently among themselves than with taxa from other modules).

**Mutualism:** symbiotic relationship that entails overall beneficial effects for both partners.

**Parasitism:** symbiotic relationship in which one partner, the parasite, benefits at the expense of the other, the host.

**Phylogenetic congruence:** given two phylogenetic trees, the extent to which each node and branch-length in one tree maps to a corresponding position in the other.

**Phylogenetic signal:** a measure of the statistical dependence among species' trait values due to their phylogenetic relationships.

**Phylogenetic tracking:** process by which the speciation of one symbiotic partner is followed by speciation of the other partner. Diversification is asynchronous.

**Pseudocospeciation:** (a.k.a. preferential host switching or host-shift speciation); symbiont diversification mimicking the tree topology of the host as a result of complete host-switching events (i.e., colonization of a new host species followed by speciation) among closely related hosts, rather than host-symbiont cospeciation. Diversification is asynchronous.

**Symbiosis:** biological interaction between two organisms of different species characterized by being long-lasting and physically close at some point if not all the time. The symbiont, or symbiote, is any member of a pair of organisms involved in this symbiotic relationship, with the larger member usually designated as the host.

**Vertical transmission:** transmission of an interaction from parents to offspring.

**Vicariance:** pattern of phylogenetic congruence caused by common biogeographic processes rather than by biological interactions between the symbiotic partners.

**Figure 1. Incorporating traits in cophylogenetic analysis.** (A) Tanglegram depicting the associations (red lines) between hosts and symbionts observed in nature. (B) Classical cophylogenetic analyses use a binary (presence/absence) matrix coding the associations. (C) The binary matrix could be replaced by a matrix of quantitative interaction traits that measure the strength of the associations. (D) If several traits are available, a set of matrices could be used. (E) Putative coevolving one-sided traits of hosts and symbionts could be measured to compute a codependence matrix expressing the agreement between the host and symbiont trait.

between individuals or taxa. One of the properties emerging from this nonrandomness is **modularity** [36]. Modules likely represent fundamental units of coevolution [23,37,38] that provide an ecologically meaningful scale for cophylogenetic assessment [11]. Searching for cophylogenetic signal within modules would facilitate linking of particular species groups to patterns stemming from their evolutionary histories.

### Linking patterns with mechanisms

Cophylogenetic signal is often interpreted as evidence of a high level of cospeciation in the system under study. However, different coevolutionary processes can lead to apparent cospeciation [16] and processes unrelated to cospeciation, such as **pseudocospeciation** can also produce some degree of phylogenetic congruence [39,40]. Althoff *et al.* [16] identified four mechanisms that account for cospeciation (Box 2), which, when acting alone or concurrently to varying extents, may induce similar levels of cophylogenetic signal. Likewise, low cophylogenetic signal does not necessarily imply the lack of an evolutionary relationship between symbiotic partners, because cophylogenetic signal can be modulated by the interactions among symbiont (and/or host) species in a community (**diffuse coevolution**) (e.g., [41,42]). For example, whereas low cophylogenetic signal has been observed between orchids and their bee pollinators, pre-existing traits in the latter (i.e., collection of aromatic compounds) likely drove floral adaptation and orchid diversification [43].

Russo *et al.* [18] recently proposed a framework that offers more mechanistic explanations to cophylogenetic patterns, thereby providing an absolute classification in which each host–symbiont system is assigned to one of eight possible outcomes (Box 2). However, we argue that **vicariance**, **phylogenetic tracking**, or coevolution are not mutually exclusive and that patterns observed in nature may well result from a combination of processes, each acting with different intensity.

In Box 2, we propose the **cophylospace**, a modified framework aimed at enhancing the explanatory power of cophylogenetic analysis, which rests on two key elements. First, measuring the strength of cophylogenetic signal becomes highly relevant. Whereas current methods (see [39,44] for reviews) provide statistical evidence for cophylogenetic signal, they produce no directly interpretable statistic as to its strength and only recently have Balbuena *et al.* [15] proposed a metric for this. Second, evaluating the correlation between the host–symbiont association matrix and the phylogeny of both symbiotic partners is crucial to gain mechanistic insights into coevolutionary processes (Box 2).

A generalized use of the cophylospace approach will require further methodological refinements. For instance, Russo *et al.* [18] used a Mantel test to assess whether phylogenetically related taxa are more likely to interact with similar partners but noted that this is not effective in specialized symbiotic associations, as often occurs with host and parasites, given that both closely and distantly related taxa can be equally dissimilar in the associations with their partners. However, Box 3 illustrates how the cophylospace approach could be applied tentatively using morphological data on the interacting partners.

### The holy trinity: genomics, heterogeneity, and time

The future of cophylogenetic analysis requires embracing, both conceptually and analytically, the ongoing spectacular progress in genomics and phylogenetic inference. Analyses of genome-wide data often result in more robust estimation of phylogenetic histories, but such phylogenomic analyses also increasingly include large numbers of taxa. Cophylogenetic analyses of large phylogenies can quickly become computationally prohibitive, in particular if the influence of



**Box 2. Cophylospace: a quantitative framework for cophylogeny**

We define a three-dimensional space (cophylospace, Figure 1) based on the processes proposed by Russo *et al.* [18]. Whereas in Russo *et al.* [18] a given cophylogenetic scenario is assigned to a single, most likely process (red crosses in Figure 1), the present setting places it in a three-dimensional space based on three quantitative parameters: cophylogenetic signal and the associations between the bipartite network of symbiotic partners and their respective phylogenies. Thus, a given cophylogenetic scenario could be explained in terms of its position along the three axes reflecting the relative contributions of the different processes (Figure 1).

High cophylogenetic signal would suggest that cospeciation is important in the system studied (Figure 1). Such cospeciation can result from four different mechanisms: coevolution, vicariance, phylogenetic tracking, and **vertical transmission** [16]. Since the latter represents a special case of phylogenetic tracking [18], it is not considered in the present framework. Assessing the strength of the association between the interaction matrix and the host and symbiont phylogenies (phylogenetic signal) could help establish the relative mechanistic contributions to the observed patterns. For instance, say phylogenetic signal between the association matrix and host phylogeny is strong and that between the symbiont phylogeny it is weak. This would be indicative of the symbiont speciation being mostly determined by that of the host (phylogenetic tracking) (Figure 1).

Lack of, or low, cophylogenetic signal does not support cospeciation as a major force (Figure 1) but does not necessarily imply that the diversification of symbiont partners is unrelated to each other. This can suggest diffuse coevolutionary scenarios instead [18]. For instance, situations in which the host–symbiont association matrix is correlated with the phylogeny of one interacting partner are indicative of asymmetric diversification [18] (Figure 1). Likewise, if the association matrix were correlated with the phylogenies of both partners, the diversification would be symmetric (Figure 1).

The cophylospace quantitative framework will facilitate comparisons between different host–symbiont systems, as well as meta-analyses to reveal general patterns (e.g., see Hayward *et al.* [71]) and provide insights into putative mechanistic explanations for those patterns (see Outstanding questions). For example, the hypothesis of widespread diffuse coevolution in plant–pollinator systems compared with host–parasite associations [18] could be tested. Similarly, we hypothesize that host phylogenetic tracking will be common in host–parasite systems.

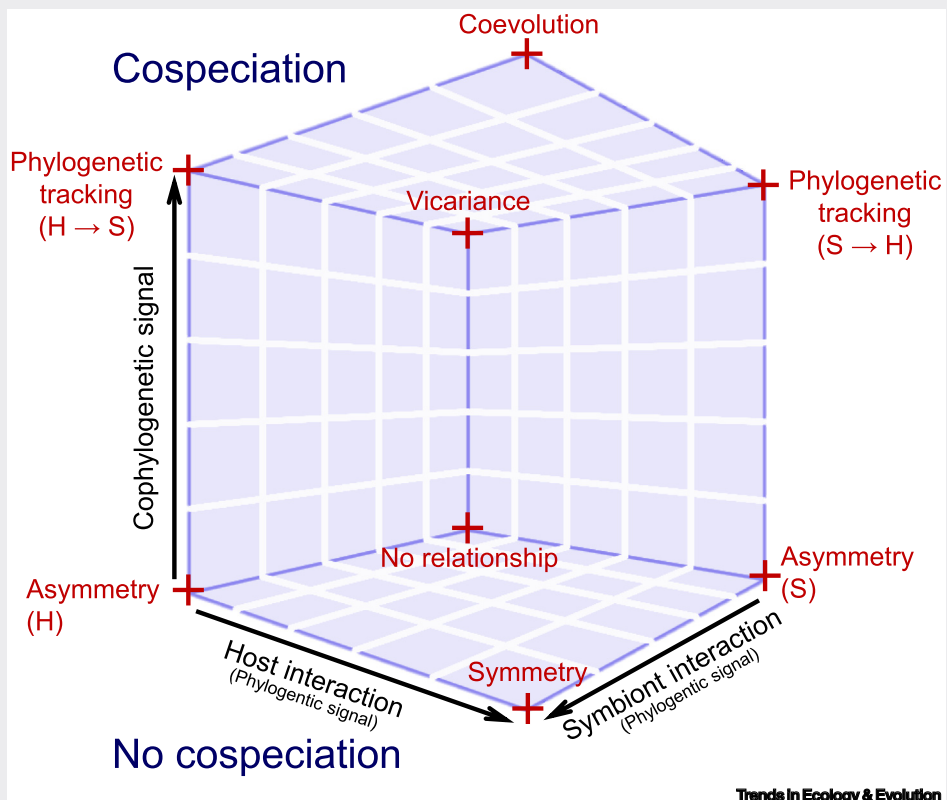


Figure 1. Cophylospace, a quantitative cophylogenetic spatial framework linking cophylogenetic patterns with underlying processes. Abbreviations: H, hosts; S, symbionts.

phylogenetic uncertainty is explicitly considered (Box 1). To help alleviate this, new methods able to accommodate actual divergence date estimates from dated phylogenies should be developed. Then, narrow estimates of dated nodes corresponding to putative codivergence events could rule out some diversification scenarios and increase the computational efficiency of event-based approaches by leaving some nodes out of the calculations. Estimates of node divergence dates might also be conveniently incorporated into hybrid events-based and global-fit methods (Box 1) to more fully capture cophylogenetic processes. Therefore, cophylogeny should aim to incorporate powerful new advances in the implementation of molecular clock methods (e.g., [45,46]).

The growing wealth and depth of available genomic data enables detailed characterizations of individual heterogeneity within populations. Current phylogenetic and phylogeographic approaches increasingly rely on assimilating such within-clade genetic heterogeneity using models that account for population genetic processes [47]. In classical cophylogeny, species are considered homogeneous entities across their distributional range and the genetic backgrounds of the interacting individuals are ignored. However, speciation involves processes acting at the microevolutionary level (e.g., shifts in allele frequencies within populations). Since ecological and evolutionary processes are mutually linked over evolutionary time scales [8,48], cophylogeny should explicitly consider intraspecific variation, both conceptually and analytically. This is particularly relevant at the community scale, where the significant units are populations rather than species (e.g., [6,35,49,50]). Furthermore, in taxa that permit it, sequencing ancestral and extant lineages within a species over time could reveal the impact of host associations on the evolutionary rate, diversification, and the genomic bases of interactions [42]. Alternatively, one could model the evolution of the host repertoire of a symbiont group [51]. Therefore, incorporating

### Box 3. Example application of the cophylospace approach

We examine a simple host–parasite system of 14 *Ligophorus* spp. (Monogenea: Dactylogyridae) on six Mediterranean and Black Sea species of grey mullets (Mugilidae) (Figure I). The R code utilized is available on Zenodo (<http://doi.org/10.5281/zenodo.4892783>). Phylogenetic information on hosts and parasites and data on their associations were obtained from Rodríguez-González *et al.* [72,73] and Durand *et al.* [74]. To evaluate whether closely related hosts are associated with more morphologically similar parasites than expected by chance, we used published information on the shape of ventral anchors, which play an important role in host attachment [72,73]. We computed a pairwise matrix of Procrustes distances between species with Procrustes coordinates of ventral anchors from 244 specimens [73]. Likewise, to test the association between parasite phylogeny and host morphology, we produced a distance matrix based on the morphometry of mullet species from [fishbase.org](http://fishbase.org). (Ideally information on gill shape should have been used but is unavailable.)

Cophylogenetic signal and interactions between phylogeny and morphology of hosts and parasites were assessed using the sum of squared residuals ( $m^2$ ) produced by PACo [64], which is inversely proportional to the concordance between the configurations tested. Significance of the associations between hosts and parasites was established based on 1000 permutations. Since  $m^2$  depends on branch length units, dissimilarity matrices were normalized prior to analysis.

Results were as follows: cophylogenetic signal (i.e., agreement between host and parasite matrices of patristic distances)  $m^2 = 0.706$ ,  $P = 0.014$ ; interaction of host phylogeny with parasite shape,  $m^2 = 0.792$ ,  $P = 0.067$ ; interaction of parasite phylogeny with host shape,  $m^2 = 0.610$ ,  $P = 0.003$  (Figure II).

Under a classical cophylogenetic analysis, the significant cophylogenetic signal detected would be interpreted as evidence of cospeciation between hosts and parasites. However, the cophylospace framework revealed that whereas phylogenetic and morphological distances of *Ligophorus* contribute similarly to explain the pattern of host–parasite associations, parasite phylogeny is more strongly associated with the morphometric traits of the hosts than with host phylogeny. The position of our system in cophylospace (Figure II) suggests some degree of asymmetry in which host traits may have influenced speciation of *Ligophorus* spp. So speciation in *Ligophorus* spp. appears to have been primarily influenced by phylogenetic tracking of host resources, which conforms to the scenario of **coadaptive codiversification** proposed by Clayton *et al.* [75] to account for host–parasite coevolution.

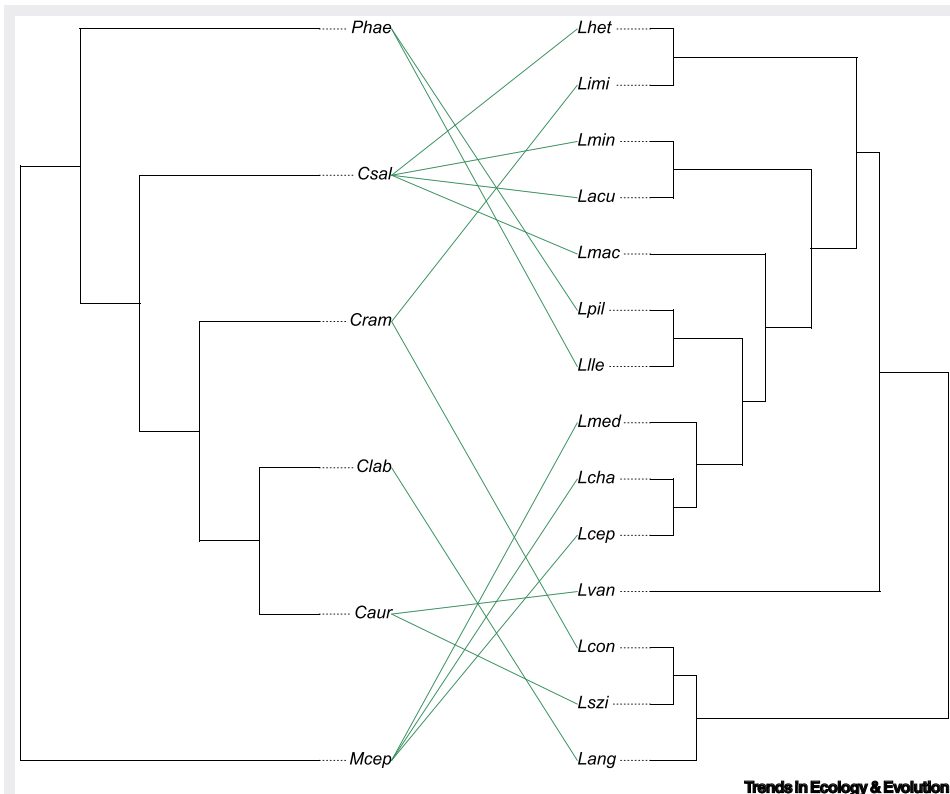


Figure 1. Tanglegram representing the association of six Mediterranean and Black Sea grey mullet species with 14 species of *Ligophorus*. Host abbreviations: *Caur*, *Chelon auratus*; *Clab*, *Chelon labrosus*; *Cram*, *Chelon ramada*; *Csal*, *Chelon saliens*; *Mcep*, *Mugil cephalus*; *Phae*, *Planiliza haematocheilus*. Parasite abbreviations: *Lacu*, *Ligophorus acuminatus*; *Lang*, *Ligophorus angustus*; *Lcep*, *Ligophorus cephalis*; *Lcha*, *Ligophorus chabaudi*; *Lcon*, *Ligophorus confusus*; *Lhet*, *Ligophorus heteronchus*; *Limi*, *Ligophorus imitans*; *Llle*, *Ligophorus llewellyni*; *Lmac*, *Ligophorus macrocolpos*; *Lmed*, *Ligophorus mediterraneus*; *Lmin*, *Ligophorus minimus*; *Lpil*, *Ligophours pilengas*; *Lszi*, *Ligophorus szidati*; *Lvan*, *Ligophorus vanbenedeni*.

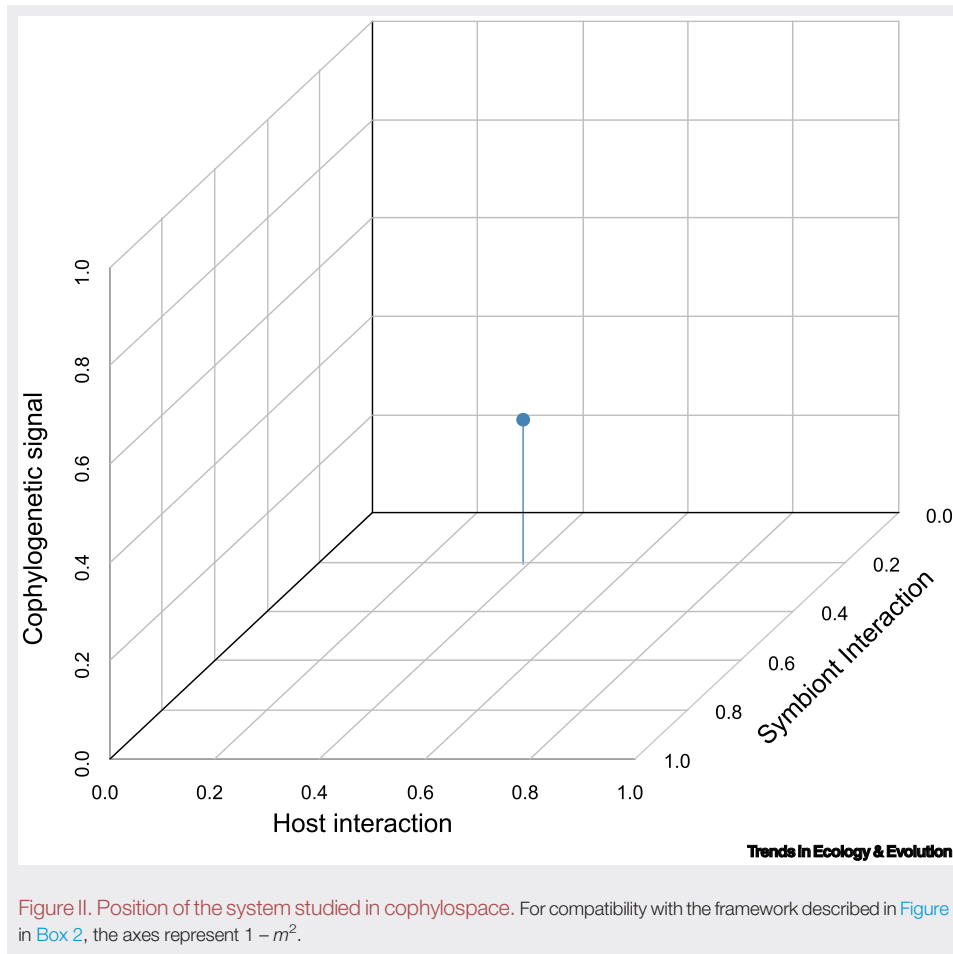
heterogeneity generated by population-level processes associated with different histories across the genome can lead to deeper insights into the processes underlying macroevolutionary change.

### Looking ahead

While there are doubtless more avenues that deserve to be explored than we cover here, three stand out to us as particularly relevant, although there is a necessity for considerable further theoretical work before they can be fully implemented.

First, the complex biology of certain organisms poses further challenges (e.g., accounting for hybridization, incomplete lineage sorting, horizontal gene transfer, etc.). In such cases, phylogenetic networks that model reticulate evolution in addition to vertical descent, can better describe the evolutionary history of a group [52]. In a phylogenetic tree, the evolutionary distance between two extant species is the path length from each species to the other via its most recent common ancestor, whereas in a phylogenetic network, there may be many paths linking two extant species [53]. Thus, new cophylogenetic methods able to accommodate this variability in evolutionary





pathways are needed (see [54]), potentially leading to novel approaches based on comparing similarities among networks instead of dichotomic trees.

Second, integration of biogeographic information in cophylogeny is highly relevant, since the geographical location of extant and ancestral nodes in the phylogenies of symbionts determines available opportunities for evolutionary events such as cospeciation and **host switching** [55]. As methods for inferring historical biogeography from phylogenetic reconstructions have greatly improved during the last two decades [56,57], robust biogeographic scenarios are becoming increasingly available. Thus, cophylogeny should progressively incorporate biogeographical hypotheses. Although the specific tools necessary for this remain scarce, recent pioneering approaches anticipate the simultaneous modeling of coevolution and biogeography [21,44,55].

Third, the majority of work to date has concentrated on exploring cophylogenetic relationships between two sets of organisms. However, shared evolutionary histories across more than two trophic levels also occur [58,59]. Considering higher-order interactions in cophylogeny can be daunting, as it represents a substantially more complex problem, because in addition to assessing the degree of overall cophylogenetic signal in the system, interactions among symbiotic partners should also be considered. For instance, in a tri-trophic system formed by X, Y, and Z, one

must tease apart the degree of phylogenetic congruence between  $X$  and  $Y$ , that is not explained by  $Z$ , from the joint phylogenetic congruence among all three partners. Unsurprisingly, the issue of assessing phylogenetic congruence among three phylogenetic trees has received considerably less attention to date (but see [60,61]) and the question of relating such congruence to evolutionary mechanisms remains to be explored.

### Concluding remarks

Cophylogeny represents a major framework to address fundamental questions in eco-evolutionary research (see [Outstanding questions](#)). However, linking patterns with mechanisms remains a significant challenge. To tackle this issue, cophylogeny must embrace, both analytically and theoretically, current hypotheses and approaches from other areas of evolutionary ecology. We stress the need to investigate the multiple scales of organization at which host–symbiont associations operate and to quantify the strength of the relationships between two-sided interaction traits and evolutionary histories.

A number of significant challenges remain. For example, how to integrate reticulate-network evolution, take full advantage of dated phylogenies, and extend the conceptual and analytical capabilities of cophylogeny to large phylogenies and across multitrophic interactions. However, several of the new directions suggested herein can be pursued at present. The cophylospace framework ([Box 2](#)) addresses several fundamental outstanding questions and can promote further meta-analyses to evaluate the generality of patterns reported to date. For instance, the cophylospace framework would allow new ways of testing the long-standing prediction that cospeciation is stronger in mutualistic systems than in host–parasite systems. Meanwhile, future availability of genomic data and incorporation of within-species heterogeneity across geographical scales can improve estimation of the evolutionary rates, diversification, and understanding of the genomic bases of symbiotic interactions, thereby illuminating the ecological processes underlying macroevolutionary change. Thus, adopting the changes proposed herein offers the transformation of cophylogenetic analysis into a powerful and more general platform to advance understanding of the evolution and diversification of interspecific interactions.

### Acknowledgments

We thank three anonymous reviewers, Prof. F. Ronquist, and the Editor, A. Stephens, for very constructive feedback and suggestions on the manuscript. We are also grateful to Dr A. Rodríguez González for assistance with datasets used in [Box 3](#). This work received funding from the Swiss National Science Foundation (SNSF grant 31003A\_169211 to I.B.-C.) and Ministry of Science and Innovation of Spain (grant PID2019-104908GB-I00 to J.A.B.). A.H. is supported by a Biotechnology and Biological Sciences Research Council (BBSRC) David Phillips Fellowship (BB/N020146/1).

### Declaration of interests

No interests are declared.

### Supplemental information

Supplemental information associated with this article can be found online at <https://doi.org/10.1016/j.tree.2021.06.006>.

### References

- Charleston, M. and Libeskind-Hadas, R. (2014) Event-based cophylogenetic comparative analysis. In *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice* (Garamszegi, L.Z., ed.), pp. 465–480, Springer
- Shapira, M. (2016) Gut microbiotas and host evolution: scaling up symbiosis. *Trends Ecol. Evol.* 31, 539–549
- Munzi, S. *et al.* (2019) When the exception becomes the rule: an integrative approach to symbiosis. *Sci. Total Environ.* 672, 855–861
- Zook, D. (2015) Symbiosis—evolution’s co-author. In *Reticulate Evolution: Symbiogenesis, Lateral Gene Transfer, Hybridization and Infectious Heredity* (Gontier, N., ed.), pp. 41–80, Springer
- Vidal, M.C. and Segraves, K.A. (2021) Coevolved mutualists experience fluctuating costs and benefits over time. *Evolution* 75, 219–230
- Lagrué, C. *et al.* (2016) Genetic structure and host–parasite co-divergence: evidence for trait-specific local adaptation. *Biol. J. Linn. Soc.* 118, 344–358

### Outstanding questions

What can we learn about the predictability and repeatability of cophylogenetic scenarios?

When do different traits suggest similar cophylogenetic scenarios? Do all traits have a similar impact on the mechanisms behind the cophylogenetic signal? Do the nuances of quantitative traits increase or decrease cophylogenetic signal?

Are cophylogenetic patterns more common in some symbiotic associations than others (e.g., mutualists versus parasitism)? Are those cophylogenetic patterns dominated by the same mechanisms in different symbiotic associations (e.g., mutualism versus parasitism)? Are there differences between strict one-to-one interactions and many-to-one and many-to-many interactions (diffuse symbioses: mycorrhiza, gut bacteria)?

Do symbionts typically codiverge with their hosts or is diversity primarily generated via host switching?

Do cophylogenetic patterns decrease over time? Does the possibility for host-switching events increase as more time elapses?

How well can cophylogenetic approaches bridge the micro- and macroevolution continuum? Do similar patterns emerge at different organizational scales?

Can cophylogenetic patterns inform us about the resilience of symbiotic relationships to global change?

7. Frickel, J. *et al.* (2018) Population size changes and selection drive patterns of parallel evolution in a host–virus system. *Nat. Commun.* 9, 1706
8. Weber, M.G. *et al.* (2017) Evolution in a community context: on integrating ecological interactions and macroevolution. *Trends Ecol. Evol.* 32, 291–304
9. Charleston, M.A. (2003) Recent results in cophylogeny mapping. In *Advances in Parasitology*, pp. 303–330, Academic Press
10. Hadfield, J.D. *et al.* (2014) A tale of two phylogenies: comparative analyses of ecological interactions. *Am. Nat.* 183, 174–187
11. Hutchinson, M.C. *et al.* (2017) Cophylogenetic signal is detectable in pollination interactions across ecological scales. *Ecology* 98, 2640–2652
12. Pollock, F.J. *et al.* (2018) Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. *Nat. Commun.* 9, 4921
13. Drinkwater, B. and Charleston, M.A. (2016) RASCAL: a randomized approach for coevolutionary analysis. *J. Comput. Biol.* 23, 218–227
14. Hutchinson, M.C. *et al.* (2017) PACo: implementing procrustean approach to cophylogeny in R. *Methods Ecol. Evol.* 8, 932–940
15. Balbuena, J.A. *et al.* (2020) Random tanglegram partitions (Random TaPas): an Alexandrian approach to the cophylogenetic Gordian knot. *Syst. Biol.* 69, 1212–1230
16. Althoff, D.M. *et al.* (2014) Testing for coevolutionary diversification: linking pattern with process. *Trends Ecol. Evol.* 29, 82–89
17. Poisot, T. (2015) When is co-phylogeny evidence of coevolution? In *Parasite Diversity and Diversification: Evolutionary Ecology Meets Phylogenetics* (Morand, S.S. *et al.*, eds), pp. 420–433, Cambridge University Press
18. Russo, L. *et al.* (2018) Quantitative evolutionary patterns in bipartite networks: vicariance, phylogenetic tracking or diffuse coevolution? *Methods Ecol. Evol.* 9, 761–772
19. Adams, D.C. and Collyer, M.L. (2019) Phylogenetic comparative methods and the evolution of multivariate phenotypes. *Annu. Rev. Ecol. Syst.* 50, 405–425
20. Delmas, E. *et al.* (2019) Analysing ecological networks of species interactions. *Biol. Rev.* 94, 16–36
21. Braga, M.P. *et al.* (2020) Bayesian inference of ancestral host–parasite interactions under a phylogenetic model of host repertoire evolution. *Syst. Biol.* 69, 1149–1162
22. Aznar, F.J. *et al.* (2001) Establishing the relative importance of sympatric definitive hosts in the transmission of the sealworm, *Pseudoterranova decipiens*: a host–community approach. *NAMMCO Sci. Pub.* 3, 161–171
23. Olesen, J.M. *et al.* (2007) The modularity of pollination networks. *Proc. Natl. Acad. Sci. U. S. A.* 104, 19891
24. Poulin, R. *et al.* (2011) Host specificity in phylogenetic and geographic space. *Trends Parasitol.* 27, 355–361
25. Nieberding, C.M. *et al.* (2010) The use of co-phylogenetic patterns to predict the nature of host–parasite interactions, and vice-versa. In *The Biogeography of Host–Parasite Interactions* (Morand, S. and Krasnov, B., eds), pp. 59–70, Oxford University Press
26. Hall, A.R. *et al.* (2020) Measuring coevolutionary dynamics in species-rich communities. *Trends Ecol. Evol.* 35, 539–550
27. Alexandersson, R. and Johnson, S.D. (2002) Pollinator-mediated selection on flower–tube length in a hawkmoth-pollinated *Gladiolus* (Iridaceae). *Proc. R. Soc. B* 269, 631–636
28. Klumpers, S.G.T. *et al.* (2019) Foraging efficiency and size matching in a plant–pollinator community: the importance of sugar content and tongue length. *Ecol. Lett.* 22, 469–479
29. Messu Mandeng, F.D. *et al.* (2015) A phylogeny of *Cichlidogyrus* spp. (Monogenea, Dactylogyridae) clarifies a host-switch between fish families and reveals an adaptive component to attachment organ morphology of this parasite genus. *Parasit. Vectors* 8, 582
30. Petersen, D.S. *et al.* (2018) Holding tight to feathers – structural specializations and attachment properties of the avian ectoparasite *Crataerina pallida* (Diptera, Hippoboscidae). *J. Exp. Biol.* 221, jeb179242
31. Manceau, M. *et al.* (2017) A unifying comparative phylogenetic framework including traits coevolving across interacting lineages. *Syst. Biol.* 66, 551–568
32. Braga, M.P. *et al.* (2018) Unifying host-associated diversification processes using butterfly–plant networks. *Nat. Commun.* 9, 1–10
33. Duron, O. *et al.* (2017) Evolutionary changes in symbiont community structure in ticks. *Mol. Ecol.* 26, 2905–2921
34. Klimov, P.B. *et al.* (2017) Detecting ancient codispersals and host shifts by double dating of host and parasite phylogenies: application in proctophylloid feather mites associated with passerine birds. *Evolution* 71, 2381–2397
35. Park, E. *et al.* (2020) Shared geographic histories and dispersal contribute to congruent phylogenies between amphipods and their microsporidian parasites at regional and global scales. *Mol. Ecol.* 29, 3330–3345
36. Fortuna, M.A. *et al.* (2010) Nestedness versus modularity in ecological networks: two sides of the same coin? *J. Anim. Ecol.* 79, 811–817
37. Bascompte, J. and Jordano, P. (2014) *Mutualistic Networks*, Princeton University Press
38. Andreazzi, C.S. *et al.* (2017) Network structure and selection asymmetry drive coevolution in species-rich antagonistic interactions. *Am. Nat.* 190, 99–115
39. de Vienne, D. *et al.* (2013) Cospeciation vs host-shift speciation: methods for testing, evidence from natural associations and relation to coevolution. *New Phytol.* 198, 347–385
40. Kahnt, B. *et al.* (2019) Should I stay or should I go? Pollinator shifts rather than cospeciation dominate the evolutionary history of South African *Rediviva* bees and their *Diascia* host plants. *Mol. Ecol.* 28, 4118–4133
41. Ashby, B. and King, K.C. (2017) Friendly foes: the evolution of host protection by a parasite. *Evol. Lett.* 1, 211–221
42. Betts, A. *et al.* (2018) High parasite diversity accelerates host adaptation and diversification. *Science* 360, 907–911
43. Ramirez, S.R. *et al.* (2011) Asynchronous diversification in a specialized plant–pollinator mutualism. *Science* 333, 1742–1746
44. Martínez-Aquino, A. (2016) Phylogenetic framework for coevolutionary studies: a compass for exploring jungles of tangled trees. *Curr. Zool.* 62, 393–403
45. Ho, S.Y.W. (2014) The changing face of the molecular evolutionary clock. *Trends Ecol. Evol.* 29, 496–503
46. Mello, B. *et al.* (2021) Molecular dating for phylogenies containing a mix of populations and species by using Bayesian and RelTime approaches. *Mol. Ecol. Resour.* 21, 122–136
47. Bravo, G.A. *et al.* (2019) Embracing heterogeneity: coalescing the Tree of Life and the future of phylogenomics. *PeerJ* 7, e6399
48. Hendry, A.P. (2019) A critique for eco-evolutionary dynamics. *Funct. Ecol.* 33, 84–94
49. Sweet, A.D. *et al.* (2018) Integrating phylogenomic and population genomic patterns in avian lice provides a more complete picture of parasite evolution. *Evolution* 72, 95–112
50. Hoyal Cuthill, J.F. and Charleston, M. (2015) Wing patterning genes and coevolution of Müllerian mimicry in *Heliconius* butterflies: support from phylogeography, cophylogeny, and divergence times. *Evolution* 69, 3082–3096
51. Braga, M.P. *et al.* (2021) Evolution of butterfly–plant networks over time, as revealed by Bayesian inference of host repertoire. *bioRxiv* Published online February 5, 2021. <https://doi.org/10.1101/2021.02.04.429735>
52. Francis, A.R. and Steel, M. (2015) Tree-like reticulation networks—When do tree-like distances also support reticulate evolution? *Math. Biosci.* 259, 12–19
53. DaCosta, J.M. and Sorenson, M.D. (2016) ddRAD-seq phylogenetics based on nucleotide, indel, and presence–absence polymorphisms: analyses of two avian genera with contrasting histories. *Mol. Phylogenet. Evol.* 94, 122–135
54. Scornavacca, C. *et al.* (2011) Tanglegrams for rooted phylogenetic trees and networks. *Bioinformatics* 27, i248–i256
55. Berry, V. *et al.* (2018) A geography-aware reconciliation method to investigate diversification patterns in host/parasite interactions. *Mol. Ecol. Resour.* 18, 1173–1184
56. Ronquist, F. and Sanmartin, I. (2011) Phylogenetic methods in biogeography. *Annu. Rev. Ecol. Syst.* 42, 441–464
57. Quintero, I. *et al.* (2015) Historical biogeography using species geographical ranges. *Syst. Biol.* 64, 1059–1073
58. Forister, M.L. and Feldman, C.R. (2011) Phylogenetic cascades and the origins of tropical diversity. *Biotropica* 43, 270–278

59. Nelson, L.A. *et al.* (2014) An emerging example of tritrophic co-evolution between flies (Diptera: Fergusoninidae) and nematodes (Nematoda: Neotylenchidae) on Myrtaceae host plants. *Biol. J. Linn. Soc.* 111, 699–718
60. Mramba, L.K. *et al.* (2013) Permutation tests for analyzing cospeciation in multiple phylogenies: applications in tri-trophic ecology. *Stat. Appl. Genet. Mol. Biol.* 12, 679–701
61. Nooney, C. *et al.* (2017) A statistical method for analysing cospeciation in tritrophic ecology using electrical circuit theory. *Stat. Appl. Genet. Mol. Biol.* 16, 313–331
62. de Vienne, D.M. *et al.* (2011) Euclidean nature of phylogenetic distance matrices. *Syst. Biol.* 60, 826–832
63. Conow, C. *et al.* (2010) Jane: a new tool for the cophylogeny reconstruction problem. *Algorithms Mol. Biol.* 5, 16
64. Balbuena, J.A. *et al.* (2013) PACo: a novel procrustes application to cophylogenetic analysis. *PLoS One* 8, e61048
65. Legendre, P. *et al.* (2002) A statistical test for host-parasite co-evolution. *Syst. Biol.* 51, 217–234
66. Nylin, S. *et al.* (2018) Embracing colonizations: a new paradigm for species association dynamics. *Trends Ecol. Evol.* 33, 4–14
67. Ronquist, F. (2002) Parsimony analysis of coevolving species associations. In *Tangled Trees: Phylogeny, Cospeciation and Coevolution* (Page, R.D.M., ed.), pp. 22–64, University of Chicago Press
68. Poulin, R. (2011) Uneven distribution of cryptic diversity among higher taxa of parasitic worms. *Biol. Lett.* 7, 241–244
69. Dallas, T. *et al.* (2017) Predicting cryptic links in host-parasite networks. *PLoS Comp. Biol.* 13, e1005557
70. Terry, J.C.D. and Lewis, O.T. (2020) Finding missing links in interaction networks. *Ecology* 101, e03047
71. Hayward, A. *et al.* (2021) A broadscale analysis of host-symbiont cophylogeny reveals the drivers of phylogenetic congruence. *Ecol. Lett.* Published online May 14, 2021. <https://doi.org/10.1111/ele.13757>
72. Rodríguez-González, A. *et al.* (2016) Evolutionary modularity and morphological integration in the haptor anchor structures of *Ligophorus* spp. (Monogenea: Dactylogyridae). *Parasitol. Res.* 115, 3519–3533
73. Rodríguez-González, A. *et al.* (2017) Evolutionary morphology in shape and size of haptor anchors in 14 *Ligophorus* spp. (Monogenea: Dactylogyridae). *PLoS One* 12, e0178367
74. Durand, J.-D. *et al.* (2012) Systematics of the grey mullets (Teleostei: Mugiliformes: Mugilidae): molecular phylogenetic evidence challenges two centuries of morphology-based taxonomy. *Mol. Phylogenet. Evol.* 64, 73–92
75. Clayton, D.H. *et al.* (2015) *Coevolution of Life on Hosts: Integrating Ecology and History*, University of Chicago Press