Darwin's entangled bank through deep time: Structural stability of mutualistic networks over evolutionary timescales

Benoît Perez-Lamarque[1], Jérémy Andréoletti[2], Baptiste Morillon[2], Orane Pion-Piola[2], Amaury Lambert[2], and Hélène Morlon[2]

[1] Université de Toulouse, Toulouse INP, CNRS, IRD, CRBE, Toulouse, France
[2] Institut de biologie de l'École normale supérieure (IBENS), École normale supérieure, CNRS, INSERM, Université PSL, 46 rue d'Ulm, 75 005 Paris, France

*corresponding author: benoit.perez-lamarque@utoulouse.fr; ORCID: 0000-0001-7112-7197

Author Contributions:
All authors designed this study. BPL performed the analyses. BPL and HM analyzed the data and wrote the first version of the manuscript and all authors contributed to the revisions.

Competing Interest Statement:
The authors declare no conflict of intersts.

Keywords: species interactions, macroevolution, bipartite network, plant mutualism, latent trait, Brownian motion.

Abstract:

Mutualistic interactions form species-rich, complex networks that play essential roles for ecosystem function. Over macroevolutionary time scales, these networks change as species emerge and go extinct, yet the stability of their structural organization remains poorly understood. Here, we show that ancestral interaction networks can be reconstructed from present-day phylogenetic and interaction data. Focusing on large biogeographic scales, we infer the ancestral networks of plant mutualisms involving arbuscular mycorrhizal fungi, bat pollinators, and bird seed dispersers. We find that these mutualistic networks exhibit a modular structure that has persisted for millions of years, maintained by the evolutionary conservatism of species interactions. As species diversify, they show limited shifts in mutualistic partners. This minimal evolutionary rewiring results in a remarkable long-term stability of mutualistic network structure.

Significance statement:

Plants have engaged in diverse mutualistic interactions since their colonization of land and the emergence of flowers and fleshy fruits. However, the macroevolution of such complex interaction networks remains poorly understood because fossil evidence of plant mutualisms is scarce. By reconstructing ancient interaction networks from present-day phylogenies and ecological data, we show that plant mutualisms with mycorrhizal fungi, pollinating bats, and seed-dispersing birds have maintained a consistent modular organization for millions of years. This persistence arises because species tend to retain their partners over deep evolutionary timescales. Our findings reveal that the architecture of mutualistic networks is remarkably stable, underscoring the enduring role of evolutionary conservatism in shaping present-day biodiversity.

Introduction:

Mutualistic interactions play a key role in ecosystems (1), and have also been central to the evolution of life on Earth (2, 3). Species involved in mutualistic interactions are interconnected through shared partners, forming complex bipartite interaction networks (4). The structure of these networks is typically non-random, often displaying nestedness – where specialist species preferentially interact with generalists, or modularity, in which distinct subsets of species interact more frequently among themselves than with species outside their group (5, 6). In some cases, both patterns can be present simultaneously (7). Historically, research has focused on local, community-scale networks, revealing that non-random structures can influence community stability. More recently, metacommunity networks compiled at regional to global scales have revealed a tendency toward modularity (8). However, this tendency is not universal (9, 10), and its underlying causes remain poorly understood.

Present-day large scale metacommunity networks result from ecological and evolutionary processes operating over millions of years (6, 11). Fundamental yet unresolved questions include when their structural organization emerged, what processes shaped it, and whether these structures are evolutionary stable (12, 13). If species distributions tend to be restricted to specific biogeographic areas, for example due to dispersal limitation or ecological filtering, we may expect the formation of stable modules aligned with biogeographic regions (8). Likewise, if mutualistic interactions are largely conserved as species diversify, we may expect the formation of stable modules composed of closely-related species (14). Conversely, if a significant fraction of species is ubiquist and mutualistic interactions are more evolutionarily labile, networks may be more nested, a structure that enhances robustness to species extinctions (5).

The degree to which species biogeographic distributions are restricted, and mutualistic interactions evolutionary conserved, is therefore expected to have important implications for network structure at large spatial scales. Species distributions depend mostly on the dispersal capacity and niche breath of species. Periods of environmental stability may favor ecological specialization, reducing niche breath and favoring modular structures. As for the phylogenetic conservatism in interaction partners, we may expect selection to favor the retention of existing beneficial associations (15). Coevolution may then reinforce reciprocal specialization and compartmentalization (16). However, several processes can counteract this

stabilizing selective force, favoring partner shift and reducing phylogenetic conservatism. These include opposing selective pressures during periods of environmental instability, competition for shared partners between closely-related species (17), pairwise coevolution across geographically distinct regions resulting in independent local adaptations (18), and the exploitation of mutualism by cheaters, which can destabilize mutualistic interactions (19). Indeed, empirical evidence for phylogenetic conservatism of mutualistic interactions is mixed (20).

Hypotheses on the emergence and evolutionary stability of network structures are particularly challenging to test. Indeed, direct observation of the fossil record is limited by both low taxonomic resolution and a lack of evidence with respect to the interactions. Only a few networks have been reconstructed from fossil deposits, and these primarily represent trophic interactions, which leave more readily interpretable evidence than mutualistic ones (21, 22). An alternative is to reconstruct ancestral networks (i.e., probable ancestral interactions) using knowledge of present-day interactions and phylogenetic relationships. Inspired by models of historical biogeography (23), Braga *et al.* (24) developed a probabilistic model for the evolution of host repertoire (*i.e.* the set of hosts a symbiont interacts with) along a symbiont phylogeny that allows reconstructing ancestral host repertoires (24, 25). However, this approach groups hosts into units corresponding to higher taxonomic levels, such as classes or families, and is computationally costly. This hampers a fine-grained characterization of ancestral network structure, particularly at large biogeographic scales.

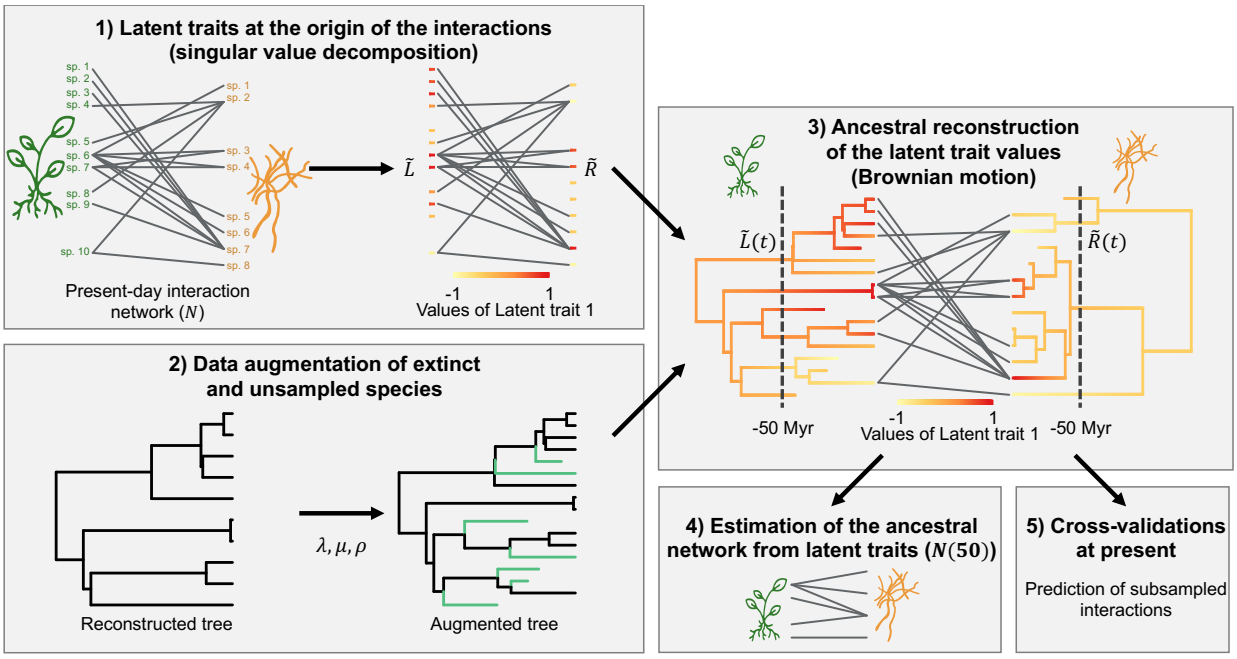
Here, we begin by developing an efficient method to reconstruct ancestral networks, based on the idea that interactions are shaped by unobserved traits evolving over the course of clades' diversification (26, 27). We then apply this approach to the large-scale mutualism of plants with arbuscular mycorrhizal fungi, bat pollinators, and bird seed dispersers. For hundreds of millions of years, plants have formed diverse and persistent interactions with animals and microorganisms, which play crucial roles in their nutrition, protection and reproduction (28–30). We reconstruct ancestral plant-mutualist networks and follow their structure through time.

Results:

We developed the "Evolution of LatEnt traits For Ancestral NeTwork reconstruction" model (ELEFANT, Figure 1, see Methods). We assume that interspecific interactions are shaped by species-specific traits, such as the production and detection of floral scents, or the morphological complementarity of spur and proboscis lengths in pollination systems. However, such traits are often unknown and/or difficult to measure in most empirical systems. Thus, instead of relying on directly measured traits, ELEFANT infers unobserved "latent" traits from the present-day interaction matrix, using singular value decomposition (26, 27), see Methods ; Figure 1.1). By construction, each species is represented by a vector of several latent trait values, and the interaction probability between a pair of species is determined by the product of their corresponding latent trait values. This interaction rule does not match classical formulations of trait matching (as, for example in (6, 31)), yet as we will show and discuss, it is effective in representing who interacts with whom and can be interpreted biologically. Next, we assume that clades evolve according to a birth-death process representing speciation and extinction events, and use data augmentation to obtain "complete" trees with probable configurations of lineages that went extinct or were not sampled (Figure 1.2). Finally, we assume that the latent traits evolve according to traditional phylogenetic models of trait evolution, such as the Brownian motion process. We use ancestral trait reconstruction to infer latent trait values along phylogenies (27, 32)(Fig. 1.3), and deduce ancestral networks (Figure 1.4). The last step (Figure 1.5) is a cross-validation procedure that assesses the ability of the model to reliably guess known interactions at present. Low scores (below 0.25) indicate that the data violates some of the model assumptions (e.g., Brownian evolution of traits), and thus that the model should not be used to infer ancestral networks.

Figure 1: General framework used to reconstruct ancestral networks

The Evolution of LatEnt traits for Ancestral Network reconsTruction framework (ELEFANT), proceeds in five steps. 1) Given a bipartite interaction network (N), a vector of d latent trait values is estimated for each extant species on each side of the network; these vectors (compiled into matrices $\tilde{L}$ and $\tilde{R}$) underlie present-day interactions (for clarity only the first latent trait value is represented here with the color code). 2) Data augmentation is used to add extinct and unsampled lineages (colored in green) on the phylogenetic trees , assuming constant rates of speciation (λ), extinction (μ), and sampling at present (ρ). 3) Ancestral values of the latent traits are estimated on the two phylogenetic trees using Brownian motion. 4) Ancestral interaction networks are reconstructed from the estimated ancestral latent trait values at any time in the past (e.g. here, 50 Mya) 5) A cross-validation procedure based on the ability to recover known interactions at present in a specific system provides a measure of confidence .



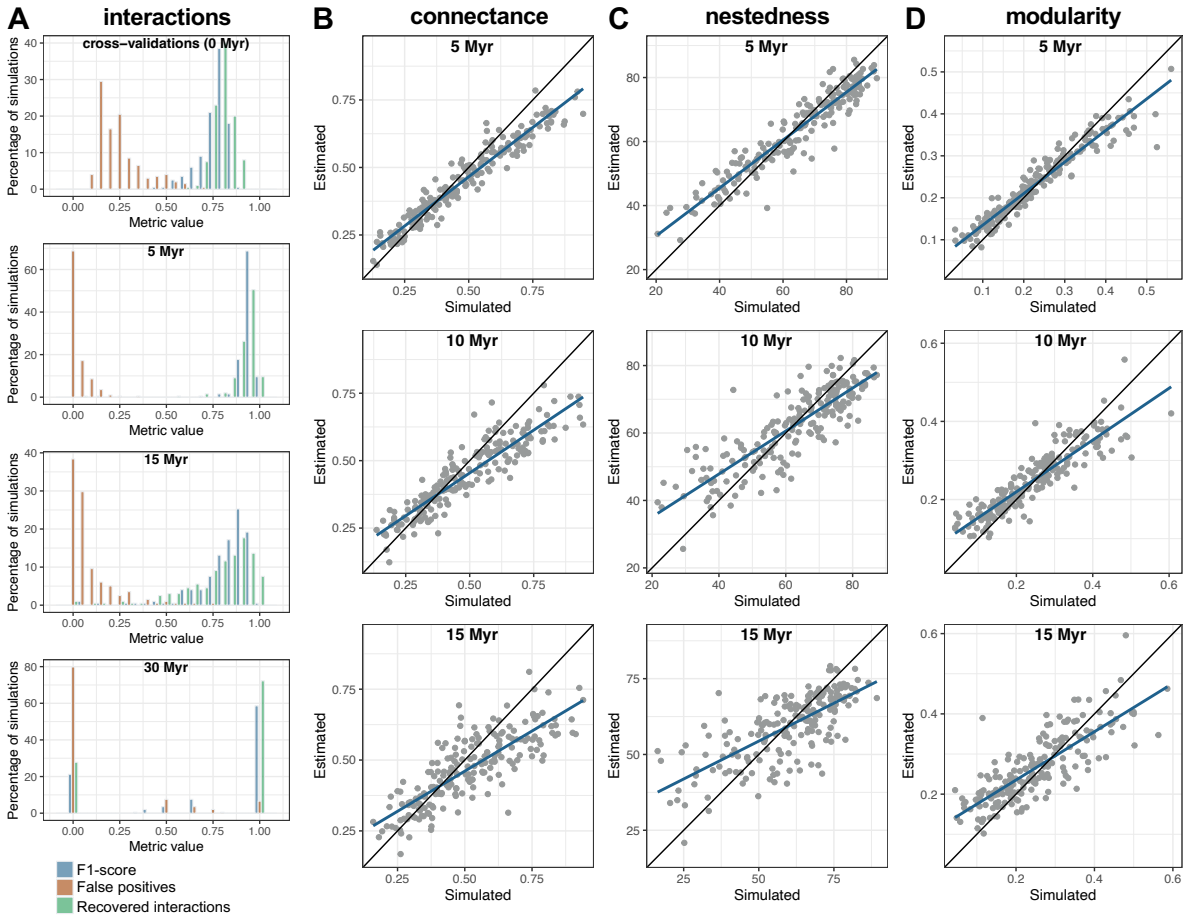
We found that ELEFANT can robustly reconstruct ancestral interaction networks as soon as interactions are not too evolutionary labile. We mimicked 30 Myr of network evolution under five distinct scenarios (see Methods): i) The first scenario considers the diversification of both hosts and symbionts, with symbionts undergoing changes in their associated hosts through "anagenetic shifts". In this process, interactions are inherited at speciation, but a symbiont lineage can change partners during its lifespan. ii) The second scenario adds the occurrence of "cladogenetic shifts", where at each symbiont speciation event, one of the daughter symbiont lineages retains its host associations, while the other forms an entirely new set of associations with randomly selected host species. iii) The third scenario corresponds to the model underlying ELEFANT, where traits evolve as Brownian processes on the host and symbiont phylogenies and their products determine whether species interact. iv) The fourth scenario, inspired from Braga *et al.* (24), models host repertoire evolution, where host repertoires are inherited at symbiont speciation, but new hosts can be acquired or lost during the lifespan of a symbiont lineage. v) Finally, we consider a scenario of "random evolution" with very frequent anagenetic shifts. In scenarios i), iii) and iv) interspecific interactions tend to be conserved over macroevolutionary timescales, while in scenarios ii) and v) they are more labile. Our cross-validation procedure efficiently discriminates these scenarios (Fig. S1), indicating that reconstructions can be trusted in scenarios i), iii), and iv), with F1-scores above 0.75 for i) and 0.5 for iii) and iv), but not in scenarios ii) and v), with F1-scores around 0.20 for ii) and below 0.10 for v). The ability to recover correct ancestral interactions (F1-scores computed from ancestral networks) is well correlated with the F1-scores of the cross-validations at present (Fig. S2): when the cross-validation indicates that ELEFANT can be trusted (F1-score >0.25), the inferred ancestral interactions are likely to be correct (F1-score >0.5). In these cases, ancestral interactions are well recovered, in particular in the recent past (Figs. 2 and S3). It follows that global metrics (connectance – *i.e.* the ratio of realized interactions –, nestedness, and modularity) and phylogenetic signal in species interactions are also well recovered, although with a tendency toward homogenization (*i.e.* the lowest and highest simulated metrics tend to be estimated closer to the average; Figs. 2, S4 & S5). The estimations of global metrics are particularly precise close to the present (5 Myr ago: mean R^2=0.79) and become noisier back in the past (20 Myr ago: mean R^2=0.32), with a good recovery of temporal trends (see Supplementary results; Fig. S6).

Figure 2: Accurate reconstruction of ancestral networks

A: Bar charts indicate the F1-score (blue), the proportion of false positives (red), and the proportion of recovered interactions (green) inferred in the cross-validations at present (upper row) and in the consensus ancestral networks (from 5 to 30 Myr ago; lower rows). 30 Mya corresponds to the most recent common ancestor of the symbionts, which interacted with a single host species. These consensus networks assemble interactions that are recovered in at least *J*% of the 'augmented ancestral networks' where *J* is the classification threshold maximizing the Youden's *J* statistic.

B-D: The estimations of the ancestral connectance (B), nestedness (C), and modularity (D) are represented at different times in the past (from 5 to 15 Myr ago) as a function of the simulated values. Each dot corresponds to one simulation, and for each simulation, we reported the estimated values averaged over 250 'augmented' ancestral networks. The black lines represent the line y=x, while the blue lines correspond to the fit of a linear model between the estimated and simulated values.

Only results obtained with the simulations of network evolution by anagenetic shifts are reported here (see Supplementary Figures for other types of simulations).



Confident in our ancestral reconstruction approach, we inferred the ancestral networks for four large plant-mutualist systems (see Methods): i) the global-scale mutualism with arbuscular mycorrhizal fungi (115 plant and 351 fungal taxa), ii) interactions with bat pollinators in the Neotropics (42 plant and 56 bat taxa), iii) bird-mediated seed dispersal in Europe (28 plant and 70 bird taxa), and iv) bird-mediated seed dispersal in the Americas (115 plant and 586 bird taxa). Cross-validation F1-scores were all above 0.25 (0.32, 0.27, 0.70, and 0.46 for i), ii), iii) and iv) respectively). Our reconstructions reveal a strikingly consistent pattern of modularity across the four systems, despite variation in species richness and composition over time (Figs. 3, 4 & 5). Plant degrees (*i.e.,* the mean number of partners per species) tend to be lower than expected under null models in three of the four systems (significantly so in two), reflecting a tendency toward specialization. Species interactions appear to be generally evolutionarily conserved (significant phylogenetic signal except in the smaller size European bird dispersal network, Figs. 3, 4 & 5). These results are robust to diversification rates and sampling fractions used to augment the trees (Fig. S7). They suggest a scenario where mutualistic associations are evolutionary conserved, favoring specialization and resulting in the formation of stable modules, where closely-related species interact with closely-related partners.

Our ancestral networks indeed reveal that angiosperms have preferentially formed associations with Glomeraceae, and non-angiosperms with early-diverging Glomeromycotina fungi, a pattern of phylogenetically structured modularity that has persisted since the origin of flowering plants (>150 Mya; Fig. 3A-C, Table S1). This tendency of flowering plants to interact with Glomeraceae ("recent plants with recent fungi") and non-flowering plants with non-Glomeraceae ("ancient plants with ancient fungi" (33)) has emerged more than 150 Myr, at the emergence of flowering plants, reinforcing the idea of an entangled diversification dynamic in these clades (34, 35). Similarly, phylogenetically structured modularity was already evident in bat pollination networks 10 Mya, corresponding to the divergence of bat subfamilies (Fig. 4). Compared with other mutualistic systems, seed dispersal networks exhibited elevated connectance (36), a pattern already present 10 Mya (Fig. 5A-B). Only a minority of taxa are specialists (Fig. 5C-D), yet the networks present a significant modularity, shaped by phylogenetically conserved functional traits that mediate frugivory (37).

Figure 3: Global scale ancestral network reconstruction of arbuscular mycorrhizal symbiosis. Consensus ancestral networks 50 (A), 100 (B), or 130 (C) Myr ago. Ancestral network metrics (D). Squares indicate a significant departure from the null model (95% central range shown in orange), whereas the crosses indicate non-significance. Error bars represent the 95% central range of values obtained from the set of augmented ancestral networks.

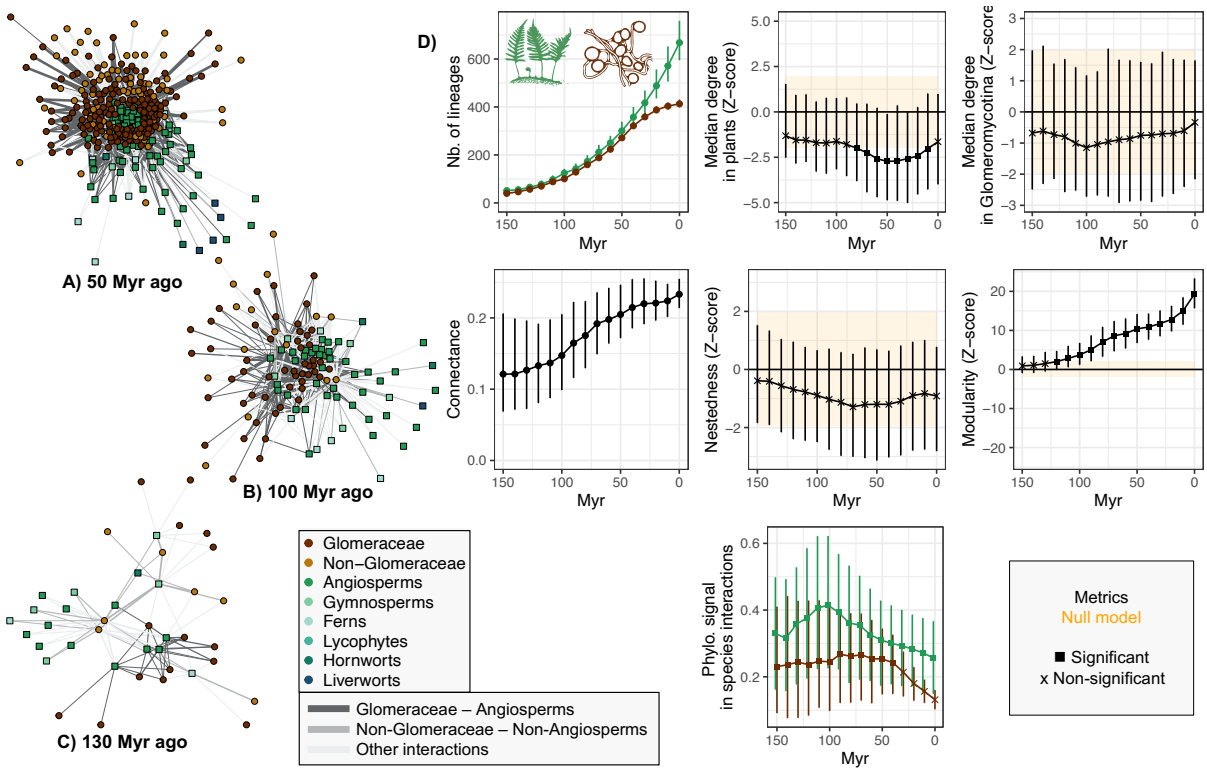


Figure 4: Ancestral network reconstruction of bat pollination in the Neotropics. Consensus ancestral network inferred 10 Myr ago (A). Ancestral network metrics (B). Squares indicate significant, and crosses non-significant departure from the null model (shown in orange), as in Fig. 3.

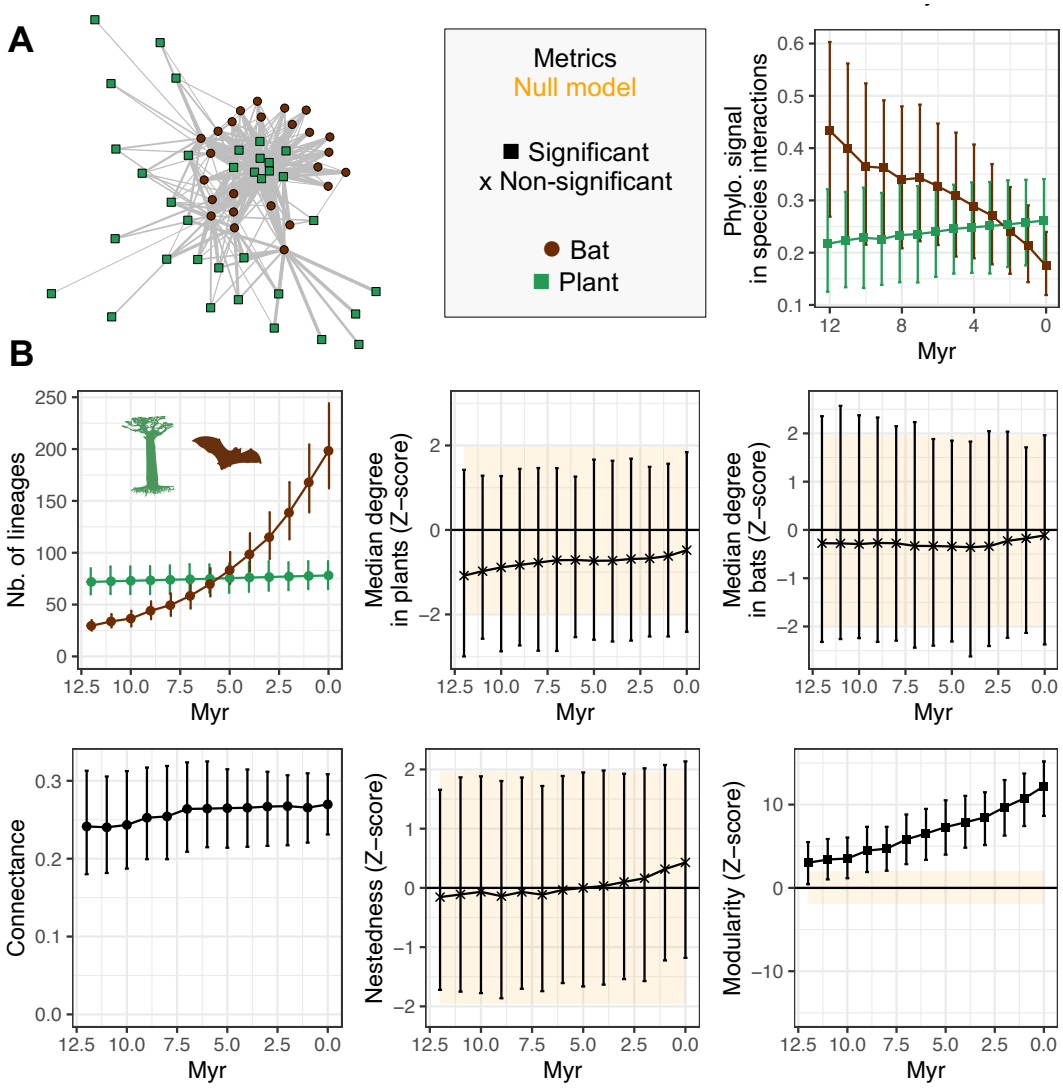
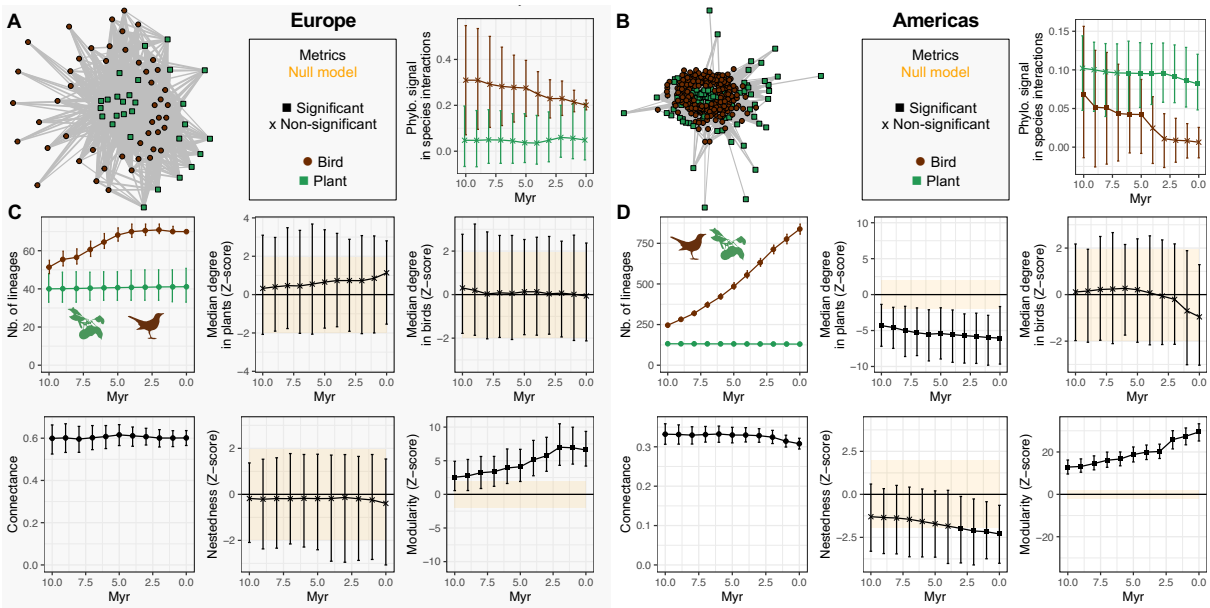


Figure 5: Ancestral network reconstruction of bird seed dispersal.

Consensus ancestral networks inferred 10 Myr ago for the European (A) and American (B) bird seed dispersal networks. Ancestral network metrics (C & D). Squares indicate significant, and crosses non-significant departure from the null model (shown in orange), as in Fig. 3 & 4.



Discussion:

Present-day metacommunity networks result from millions of years of coevolution among interacting species, but we lack direct observations of their deep-time evolution. We addressed this issue by developing a phylogenetic approach to reconstruct ancestral interaction networks. Ancestral networks of plant mutualisms involving arbuscular mycorrhizal fungi, bat pollinators, and bird seed dispersers show a consistent pattern of phylogenetically structured modularity that has persisted over macroevolutionary time scales.

The main idea underlying our approach, in comparison with previous attempts (24), is to introduce latent traits supposed to determine who interacts with whom, and to reconstruct these latent traits rather than directly the interactions. This is justified by the fact that interspecific interactions generally require a minimal level of compatibility between morphological, physiological, phenological, and/or behavioral traits to establish (1, 38). Hence, ideally, latent traits reflect biological traits that shape biotic interactions (4, 26). At the scales considered in the paper, latent traits may also reflect biogeography, which can constrain co-occurrence, and therefore potential interactions (39, 40). We used a singular value decomposition of the present-day interaction matrix to estimate latent traits, following (26, 27). By construction, these latent traits are independent, determine different axes of variations in present-day interactions, and two species are more likely to interact if they exhibit high values for the same latent traits. While this interaction rule does not match traditional formulations of trait matching or phenotypic differences, one biological interpretation is to view each latent trait as the resulting expression of a set of pleiotropic genes involved in the establishment of the interactions: if these complementary sets of genes are highly expressed in both species, the interaction can establish. Regardless of how the latent traits from the singular value decomposition are biologically interpreted – or of the specific underlying interaction rule—we showed using simulations that the approach reliably recovers ancestral interactions across diverse evolutionary scenarios, demonstrating the flexibility of latent traits in capturing who interacts with whom.

The introduction of latent traits greatly simplifies the ancestral network reconstruction problem, as ancestral trait values can easily be obtained on the host and symbiont phylogenies under classical models of trait evolution. In addition, the independency of the traits guaranteed by the singular value decomposition means that traits can be

assumed to evolve independently. We used here independent Brownian motion processes, a general framework for trait evolution (41), but other models of trait evolution, such as the Early Burst or Ornstein-Uhlenbeck processes could be used if more appropriate. The use of the Brownian motion, which is prone to smoothing, may explain the observed tendency towards homogenization of reconstructed ancestral networks. Future developments could explicitly account for biogeographic range evolution and consider only interactions between species present in the same region (as, for example, in (42)).

Another strength of our approach is to explicitly account for extinct and unsampled lineages. While we found that our results are robust when not augmenting the trees, neglecting these missing lineages results in a large undersampling of past diversity, and could affect measures of ancestral network structure (43). We used here a constant-rate homogeneous birth-death model to complement the reconstructed phylogeny, but our framework could take advantage of recent developments that provide 'augmented trees' under more realistic diversification models with heterogenous speciation and extinction rates through time and across lineages (44, 45).

Despite the flexibility of our approach, its performances vary: It performs best when modifications in the interactions occur anagenetically rather than cladogenetically, and when interactions are evolutionary conserved. When frequent interaction shifts occur, ancestral species are often inferred to interact while they do not (Fig. S3). Fortunately, our cross-validation procedure allows to assess the reliability of the ancestral reconstructions for any given empirical system. If the cross-validation F1-score exceeds 0.25, the reconstructions can be trusted. The reliability of the reconstruction also varies with time: as in all ancestral reconstruction techniques, the estimates are less precise as we go deeper into the past. As a rule of thumb, we would not consider ancestral networks constructed past halfway through the clade's age to be reliable.

The ancestral plant-mycorrhizal, pollination, and seed dispersal networks reveal a remarkably stable modular structure in the last millions of years, despite continual changes in species composition driven by speciation and extinction, and variations in environmental conditions. This modularity is phylogenetically structured, with closely related species interacting with subsets of closely-related partners. Such cophylogenetic signal (*sensu* (39)) has been reported previously (38), but we show here that it is an evolutionarily stable characteristic of these mutualistic networks, rather

than a feature that emerged only recently. These results suggest that selection favors the retention of existing beneficial associations over macroevolutionary time scales (15), even if biogeographic barriers may also play a role. This in turn favors coevolution, reinforcing reciprocal specialization and compartmentalization (16). Our findings provide further support for the stability of mutualisms across macroevolutionary timescales associated with niche conservatism in species interactions (13, 20, 29). Such evolutionary stability of mutualistic network structure over millions of years now stands in sharp contrast to its rapid erosion under human pressures, which are driving an unprecedented, worldwide homogenization of ecological communities (8, 46).

By reconstructing ancestral networks, we can follow Darwin's 'entangled bank' through deep time. Tracing back the mutualistic side of this entangled bank suggests that the benefits of retaining established mutualistic relationships outweigh its costs, resulting in stable modules of closely-related species. Similar analyses of the antagonistic side may reveal much more instability, with Red Queen dynamics driving frequent partner shifts. This deep-time perspective on interspecific interactions lays the groundwork for understanding their role in species diversification – that is, how the entangled bank modulates global biodiversity patterns.

Methods:

The ELEFANT model

We consider two interacting guilds composed of h host species and p symbiont species sampled at present. The binary matrix $N\ (h, p)$ indicates whether each pair of host and symbiont species interact (1) or not (0). The extant, sampled species of each guild have an associated reconstructed, time-calibrated phylogeny. We developed a framework, the "Evolution of LatEnt traits For Ancestral NeTwork reconstruction" (ELEFANT), illustrated in Fig. 1 and detailed in the next paragraphs, that infers ancestral network from these combined network and phylogenetic data. The framework proceeds in five steps, described below.

i) Inferring present-day latent traits:

First, assuming that interactions are shaped by unobserved traits, we obtain the theoretical values of these latent traits for each extant species using the random dot product graph (RDPG) method (26, 27)(Fig. 1.1). In previous developments, the approach was used to infer present-day interactions for species which interacting partners were not known by inputting latent trait values to extant species, *i.e.* phylogenetic tips (27, 32). Here, we aim to infer ancestral interactions. This method performs a singular value decomposition of the present-day interaction matrix N by finding the matrix $\Sigma\ (h, p)$ with null non-diagonal elements, and the matrices $L\ (h, h)$ and $R\ (p, p)$, such that $N = L \times \Sigma \times R^t$. Each diagonal element of the Σ matrix corresponds to a singular value associated with a given share of the variance observed in present-day interactions. By construction, these singular values are independent and each of them defines an independent axis of variation of a latent trait. We selected the d first eigenvalues $(\Sigma|d)$ that explain at least 90% of the total variance observed in present-day interactions. The values of the d independent latent traits associated to the h hosts are given by the $h \times d$ matrix $\tilde{L} = (L|d)$, and those associated to the p symbionts are given by the $d \times p$ matrix $\tilde{R} = (R|d)^t$. Each species is thus characterized by a vector of d latent trait values. The interaction matrix N is approximated by $\widetilde{L} \times \Sigma|d \times \tilde{R}$. Thus, the sum of the products of each latent trait between host i and symbiont j, $\widetilde{L}_i \times \Sigma|d \times \widetilde{R}_j$, reflects their likelihood of interaction. In other words, the probability of interaction between two species increases when both species exhibit high values for the same latent traits. Consequently, specialist species tend to have on average lower latent trait values, whereas generalist species tend to have higher latent trait values.

To determine the threshold J above which we can consider that there is an interaction between two species, we use a binary classification approach between the observed interaction network (N) and its approximation ($\widetilde{L} \times \Sigma | d \times \tilde{R}$) (47): we take the J value maximizing the Youden's J statistic, defined as J = sensitivity (percentage of interactions correctly identified) + specificity (percentage of non-interactions correctly identified) – 1. This value is thus specific to each network. Other thresholds yielded similar results, with Youden's J threshold providing a good compromise (Fig. S3).

ii) Augmenting phylogenetic trees:

The second step (Fig. 1.2) is required to account for lineages that were likely involved in ancestral interactions but are missing in reconstructed trees, which lack all the lineages that left no descendants at present, either because they went extinct or were not sampled. For simplicity, we consider that species diversification in each clade follows a constant-rate, homogeneous birth-death process, and we implement a data augmentation approach to simulate extinct and unsampled lineages on the host and symbiont phylogenetic trees. Given the speciation rate λ, the extinction rate μ, and the sampling fraction at present ρ, all three assumed to be constant and known (in practice, λ and μ are estimated by maximum likelihood), we augment the phylogenetic trees with extinct and unsampled lineages (see Supplementary Methods). We account for stochasticity by independently replicating this data augmentation step 250 times. In the following steps, we thus work with a set of 'augmented' phylogenetic trees for each clade, which represent probable diversification histories.

iii) Inferring ancestral latent traits:

Next, for each pair of host and symbiont augmented trees, we perform ancestral state reconstructions to estimate the latent trait values of each lineage back in time (Fig. 1.3). We assume that the d latent traits of the hosts ($\widetilde{L}$) and their symbionts ($\tilde{R}$) evolved according to independent Brownian motion processes. For each latent trait, we infer the parameters of the Brownian motion process using maximum likelihood (function mvBM in the mvMORPH R-package (48)). Then, at each time point on the augmented phylogenetic trees, we determine the expected value of each latent trait and their variances (49, 50) and perform 250 stochastic character mappings.

iv) Inferring ancestral networks:

In the fourth step, we reconstruct the ancestral network $N(t)$ between coexisting species at any given time t in the past (Fig. 1.4), based on the likelihood that two species interact given their inferred latent trait values $N(t) \sim \widetilde{L}(t) \times \Sigma|d \times \tilde{R}(t)$ and according to Youden's J threshold (see above). In the case of obligatory interactions, where a species cannot survive without at least one partner, we force every species to form at least one ancestral interaction: if the Youden's statistic is not reached with any partner, we choose the partner species with the highest likelihood of interaction. This only happens in a limited number of cases (<2% of the species in the empirical applications).

v) Cross-validations:

The last step of ELEFANT aims to provide a measure of confidence in the ancestral network reconstructions, given the strong assumptions made throughout: i) interactions determined by latent traits derived from a truncated singular value decomposition at present, ii) latent traits evolving according to independent Brownian motion processes, and iii) latent traits, and therefore interactions, inherited at speciation. We use cross-validation on the observed present-day interaction data to obtain a measure of confidence in our model predictive accuracy (Fig. 1.5): we randomly subsample 90% of the extant species, predict the partners of the other 10%, repeat this procedure 100 times, and compute the average F1-score, defined as the harmonic mean of the sensitivity (percentage of interactions correctly recovered) and the precision (the ratio of recovered interactions corresponding to true interactions). Based on the results obtained with simulations (Fig. S2), we consider reconstructions reliable when the F1-score exceeds 0.25. Although this threshold may appear low, it reflects the intrinsic difficulty of our cross-validation design, in which ELEFANT must reconstruct ancestral interactions from present-day sampled species and subsequently predict present-day interactions of unsampled species from those ancestral reconstructions – that is, present to past to present. By contrast, the primary application of ELEFANT involves the reconstruction of ancestral interactions from present-day data (present to past), a task that consistently yields substantially higher F1-scores and thus greater reliability (see simulation results in Fig. 2D and Supplementary Figures).

Taking as input an interaction matrix (the network) and associated time-calibrated phylogenies, ELEFANT thus outputs at each time *t* in the past a set of 'augmented' ancestral networks representing probable interactions between both lineages present in the phylogeny and augmented lineages, as well as a measure of confidence in these reconstructions. When the two phylogenetic trees have different ages, ELEFANT provides reconstructions until the most recent common ancestor of the youngest clade.

Summarizing ancestral networks:

We summarized the structure of augmented ancestral networks by reporting several global metrics: i) species degrees, *i.e.* the average number of symbionts interacting with each host and *vice-versa*, ii) connectance (the ratio of realized interactions), iii) nestedness (the tendency of specialist species to interact with generalist partners, measured using the NODF2 index (51)), iv) modularity (the percentage of interactions belonging to subsets of species preferentially interacting together – computed using the function computeModules (51)), and v) the phylogenetic signal of species interactions in each guild, *i.e.,* the tendency of closely related species to interact with similar partners, measured using Mantel tests (52).

We also generated consensus ancestral networks constituted of frequently recovered interactions between pairs of species (25). We added an interaction in the consensus network if it was recovered in at least x% of the augmented ancestral networks, with x being either 10%, 25%, 50%, or *J*% (Youden's *J* threshold). Consensus networks do not contain interactions involving augmented species, are these vary across augmented trees.

Simulations

We evaluated the ability of ELEFANT to reconstruct ancestral networks thanks to five simulated scenarios. We designed them based on different assumptions on how interaction networks could evolve and tried to obtain realistic simulated networks similar to empirical ones.

i) Network evolution by anagenetic shifts:

We simulated the evolution of host-symbiont interaction networks where the symbionts are obligately associated with at least one host species, while the hosts are

only facultatively interacting with symbionts. At first, we considered that the most recent common ancestor of the symbionts interacted with one single host lineage picked at random among the host species that started diversifying 10 million years before the symbionts. Then, we assumed that interactions can change during the evolution of the lineages (anagenetic shifts of interactions) and are inherited at speciation events (no cladogenetic changes). The set of associated host species of each symbiont lineage evolves independently through time according to a continuous-time Markov process: a symbiont lineage can acquire a new interaction with another host species at a rate of 0.015 or lose one of its associated host species at a rate of 0.015. Yet, a symbiont lineage cannot lose its only interaction. When a symbiont lineage speciates, the two new symbiont lineages inherit the same host partners. When a host lineage speciates, each symbiont species interacting with that host begins interacting with the two new host lineages. Finally, when a host lineage goes extinct, its associated symbionts lose this partner. If this was their only interaction, a new interaction is randomly assigned from among the surviving host lineages.

ii) Network evolution by cladogenetic shifts:

By modeling the evolution of latent traits using Brownian motions, ELEFANT assumes that the latent trait values are inherited at speciation and therefore that the two daughter species are likely to keep the same interactions. However, there are examples of closely related species that likely switched partners at speciation (53, 54) and such cladogenetic shifts may not be well handled by ELEFANT. We therefore simulated a scenario where network evolution happened through both anagenetic and cladogenetic shifts. We built upon the scenario "network evolution by anagenetic shifts" with a rate of anagenetic acquisition of 0.15 and a rate of anagenetic loss of 0.05 and added cladogenetic shifts. At each symbiont speciation event, one of the daughter symbiont lineages retains the same associated host species, while the other daughter lineage loses all its association(s) and forms a new association with a randomly selected host species.

iii) Evolution of latent traits by Brownian motions:

Next, we considered one simulated scenario corresponding to the underlying model of ELEFANT. We simulated the independent evolution of d=10 latent traits using Brownian motions on the host and symbiont phylogenetic trees with 0 as ancestral values and a variance of 0.01. We attributed interactions between pairs of host and

549 symbiont species if the product of their latent traits was larger than 0.25. Both host and
550 symbiont species were considered to be obligatory partners.
551
552 **iv) Host repertoire evolution:**
553
554 In this scenario, we assumed that host taxa diversified before the most recent common
555 ancestor of the symbionts (*e.g.* it can correspond to higher taxonomic levels, *e.g.* host
556 families) and we tracked the set of host taxa associated with each symbiont species
557 during their diversification. Following Braga *et al.* (24), we referred to this set of
558 associated host taxa as the host repertoire. Each symbiont obligatorily interacts with
559 at least one host taxon, while the hosts only facultatively interact with symbionts. We
560 started the simulation by randomly attributing one host taxon to the most recent
561 common ancestor of the symbionts. Then, the host repertoire of each symbiont lineage
562 evolved independently by acquiring new host taxa at a rate of 0.15 and losing host taxa
563 at a rate of 0.05 following a continuous-time Markov process. Host repertoires are
564 inherited at symbiont speciation.
565
566 **v) Random evolution:**
567
568 Finally, we considered a scenario of fast interaction evolution where network
569 evolution occur by anagenetic shifts at very high rates, such that interspecific
570 interactions are not conserved over macroevolutionary timescales. To simulate this
571 scenario, we used the first simulation scenario with a rate of host acquisition of 100
572 and a rate of host loss of 200.
573
574 In each of the five scenarios, we simulated network evolution for 30 million years and
575 recorded the simulated ancestral networks every 5 million years. Both the hosts and
576 the symbionts diversified according to a constant rate birth-death process with a
577 speciation rate (λ) uniformly sampled between 0.05 and 0.2 and an extinction rate (μ)
578 between 0.01 and 0.05. For computational reasons, we only kept simulated systems
579 having between 25 and 150 extant species per guild. We generated a total of 200
580 simulations per scenario.
581
582 We assumed that not all extant species were sampled at present and used a sampling
583 probability (ρ) uniformly sampled between 60% and 100%, as undersampling is a
584 common phenomenon in empirical networks (43). We discarded extinct and
585 unsampled species from the phylogenetic trees of each guild; we referred to these new

trees as the 'reconstructed trees'. In addition, we considered that not all interactions are sampled at present and that present-day interactions between sampled species have only a 90% chance of being observed while keeping at least one observed interaction per species. We referred to these subsampled present-day networks as the 'observed networks'.

The values of the parameters used in the different simulations were chosen to obtain a range of realistic simulated networks mimicking the variability observed in empirical ones. Indeed, we obtained ranges of size, connectance, modularity, and nestedness values comparable to those of empirical networks (Fig. S4)(36). We also obtained a range of phylogenetic signals in species interactions (Supplementary Results, Fig. S8), with many networks showing low signals as often observed in nature (55).

We then ran ELEFANT on each simulation composed of one observed network and two reconstructed trees. For the step of data augmentation of extinct and unsampled lineages, we used the simulated parameters values λ, μ, and ρ, and generated a total of 250 augmented ancestral networks. Between 5 and 15 singular values were often necessary to explain at least 90% of the present-day interactions in the different types of simulated networks (Supplementary Results; Fig. S9).

We then compared the simulated *versus* inferred global metrics of the ancestral networks (number of species, connectance, nestedness, and modularity) by fitting linear models. We also evaluated whether the analyses recovered significant phylogenetic signals when present in the simulations. Finally, we looked at the inferred ancestral interactions present in the consensus ancestral networks and reported the sensitivity (percentage of interactions correctly identified), the percentage of interactions corresponding to false positives, and the corresponding F1-score. We replicated the analyses without data augmentation (μ=0 and ρ=1) and confirmed that adding extinct and unsampled lineages significantly increased the performance of ELEFANT (Supplementary Results; Fig. S10).

The implementation of ELEFANT and all simulations were performed in R (56).

Empirical applications

We applied ELEFANT to three different mutualistic systems involving plants and covering a broad range of temporal and spatial scales: the global-scale arbuscular

mycorrhizal symbiosis, the pollination by bats in the Neotropics, and the seed dispersal by birds in Europe and the Americas. Although we presented ELEFANT as a species-level model applied to monophyletic clades, it can technically be applied at any taxonomic level (*e.g.* genus or family), and to non-monophyletic groups of organisms, as long as the taxa included in the network and in the phylogenetic trees match. For instance, if the phylogenetic signal and the F1-score of the cross-validations are too scarce at the species level, it is possible to merge the lineages into higher-level taxa (*e.g.* genera or families) until reaching enough signal. Here, we considered plants at the family level and plant mutualists correspond ecologically- and/or geographically-defined groups of taxa (*e.g.* nectar feeding bats, bird seed dispersers in Europe). In this case, the "augmentation" step of the framework should be considered as an imperfect attempt to account for missing lineages, rather than as a rigorous reconstruction of past diversification and diversity histories (as, in, *e.g.* (44, 45)). For example, the birth-death model is not ideal to represent the (dis)appearance of higher taxa (see, *e.g.* (57) for a better model), and only makes sense in the case of ecologically- and/or geographically-defined groups of taxa for which it can be assumed that ecological and/or geographical transitions have been rare in the time period over which networks are reconstructed. Given these imperfections, we assess the robustness of our results to choices made during the data augmentation step.

First, we reconstructed ancestral interaction networks of the arbuscular mycorrhizal symbiosis between land plants and Glomeromycotina fungi at the global scale. We gathered the interaction data from the MaarjAM database ((58); https://maarjam.ut.ee) filtered according to (10). We considered the network at the levels of virtual taxa (VT) for the Glomeromycotina (391 taxa delimited using DNA-based criteria; (58)) and plant families (115 families). We used a phylogenetic tree of Glomeromycotina from (34) and reconstructed the plant phylogeny by pruning the mega-tree of vascular plants using V-PhyloMaker2 (59). We ran ELEFANT with λ=0.012, μ=0.001, and ρ=0.86 for Glomeromycotina following the estimates of (34), and λ=0.02, μ=0.003, and ρ=0.19 for the plants (the two former parameters were estimated by fitting a constant-rate birth-death model in RPANDA on the family-level phylogeny of land plants (60) with a sampling fraction computed as the number of sampled plant families (115) divided by the estimated total number of plant families (600)). We looked at the ancestral networks for the last 150 Myr.

Second, we investigated the evolution of plant-bat pollination networks. We gathered the interaction data from (61). We focused our analysis on the evolution of nectar-

660 feeding bat pollinators from the Phyllostomidae family, a Neotropical bat family
661 including many nectar-feeding species. We obtained the phylogenetic tree of the 56
662 bat species from (62) and used V-PhyloMaker2 for the phylogenetic tree of the 42 plant
663 families. We ran ELEFANT with λ=0.20, μ=0.01 (62), and ρ=0.31 for the bats (obtained
664 by assuming that there are 180 extant species of the nectar-feeding Phyllostomidae
665 family (63)) and λ=0.02, μ=0.003, and ρ=0.63 for the bat-pollinated plants (64). We
666 looked at the ancestral networks for the last 12 Myr.
667
668 Third, we studied seed dispersal of fleshy fruits by birds in Europe and the Americas.
669 We obtained the plant-bird interactions from (8), the bird phylogenetic tree from (65),
670 and the plant phylogenetic tree from V-PhyloMaker2. Europe and Americas were
671 picked as they formed two well-sampled, separated modules in the global-scale
672 network of seed dispersal by birds (8). In Europe, the network contained 70 bird
673 species and 28 plant families forming fleshy fruits, with sampling fractions (ρ) of 1 and
674 0.78, respectively (66, 67). In the Americas, it contained 586 bird species and 115 plant
675 families forming fleshy fruits, with sampling fractions (ρ) of 0.73 and 0.93, respectively
676 (66). We ran ELEFANT with λ=0.15 and μ=0.01 for the birds (68) and λ=0.02 and μ=0.003
677 for the plants. We looked at the ancestral networks for the last 10 Myr.
678
679 The three empirical networks showed different levels of phylogenetic signal. In the
680 arbuscular mycorrhizal symbiosis, significant signal was detected in both plants
681 (R=0.30, p<0.001) and fungi (R=0.11, p<0.001). Similarly, the pollination network
682 exhibits significant phylogenetic signal in both plants (R=0.13, p=0.01) and bats
683 (R=0.12, p<0.001). In contrast, in the European seed dispersal network, closely related
684 birds interact with similar plant families (R=0.21, p<0.01), but not the reverse (R=-0.01,
685 p=0.50), and no significant signal was observed in the American seed dispersal
686 network (p-values>0.05). In these mutualistic networks, respectively, 38, 12, 4, and 18
687 latent traits were necessary to explain at least 90% of the present-day interactions.
688
689 To evaluate the significance of the structural properties of the ancestral networks
690 (nestedness, modularity, and phylogenetic signals), we compared the original metrics
691 with that obtained under a null model. Following (25), for each augmented ancestral
692 network, we used the "r00" null model (function *nullmodel* in the R-package vegan),
693 which randomizes the interactions while maintaining the network connectance. To
694 evaluate the robustness of our results, we reran ELEFANT without using data
695 augmentation or with other rates of diversification for the step of data augmentation
696 (Table S2; Fig. S7).

Acknowledgments:

The authors acknowledge G. Dalla Riva for helpful discussions of the RDPG methods as well as the members of the 'Modeling Biodiversity' team at IBENS for constructive feedbacks.
This work was performed using HPC resources from GENCI-IDRIS (Grant 2022-AD010313735).

Data availability:

All the functions developed for running ELEFANT are available on GitHub: https://github.com/BPerezLamarque/ELEFANT
All scripts and empirical data used in this manuscript will be available on OSF upon publication.

References:

1. J. L. Bronstein, *Mutualism*, J. L. Bronstein, Ed. (Oxford University Press, 2015) https:/doi.org/10.1093/acprof:oso/9780199675654.001.0001.
2. D. H. Hembry, M. G. Weber, Ecological interactions and macroevolution: A new field with old roots. *Annu. Rev. Ecol. Evol. Syst.* **51**, 215–243 (2020).
3. L. J. Harmon, *et al.*, Detecting the macroevolutionary signal of species interactions. *J. Evol. Biol.* **32**, 769–782 (2019).
4. J. Bascompte, P. Jordano, *Mutualistic networks* (Princeton University Press, 2013).
5. E. Thébault, C. Fontaine, Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science (80-.).* **329**, 853–856 (2010).
6. O. Maliet, N. Loeuille, H. Morlon, An individual based model for the eco-evolutionary emergence of bipartite interaction networks. *Ecol. Lett.* **23**, ele.13592 (2020).
7. M. A. Fortuna, *et al.*, Nestedness versus modularity in ecological networks: Two sides of the same coin? *J. Anim. Ecol.* **79**, 811–817 (2010).
8. E. C. Fricke, J. Svenning, Accelerating homogenization of the global plant – frugivore meta-network. *Nature* **585** (2020).
9. J. Davison, *et al.*, Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism. *Science (80-.).* **349**, 970–973 (2015).

10. B. Perez-Lamarque, M. A. Selosse, M. Öpik, H. Morlon, F. Martos, Cheating in arbuscular mycorrhizal mutualism: a network and phylogenetic analysis of mycoheterotrophy. *New Phytol.* **226**, 1822–1835 (2020).
11. T. Poisot, D. Stouffer, How ecological networks evolve. *bioRxiv*, 071993 (2016).
12. S. T. Segar, *et al.*, The role of evolution in shaping ecological networks. *Trends Ecol. Evol.* **35**, 454–466 (2020).
13. G. Burin, P. R. Guimarães, T. B. Quental, Macroevolutionary stability predicts interaction patterns of species in seed dispersal networks. *Science (80-.).* **372**, 733–737 (2021).
14. M. P. Braga, P. R. Guimarães, C. W. Wheat, S. Nylin, N. Janz, Unifying host-associated diversification processes using butterfly–plant networks. *Nat. Commun.* **9**, 5155 (2018).
15. P. R. Guimarães, P. Jordano, J. N. Thompson, Evolution and coevolution in mutualistic networks. *Ecol. Lett.* **14**, 877–885 (2011).
16. J. M. Olesen, J. Bascompte, Y. L. Dupont, P. Jordano, The modularity of pollination networks. *Proc. Natl. Acad. Sci.* **104**, 19891–19896 (2007).
17. M. L. Stanton, Interacting guilds: Moving beyond the pairwise perspective on mutualisms in *American Naturalist*, (2003), pp. S10–S23.
18. J. N. Thompson, B. M. Cunningham, Geographic structure and dynamics of coevolutionary selection. *Nature* **417**, 735–738 (2002).
19. J. Genini, L. P. C. Morellato, P. R. Guimarães, J. M. Olesen, Cheaters in mutualism networks. *Biol. Lett.* **6**, 494–497 (2010).
20. J. M. Gómez, M. Verdú, F. Perfectti, Ecological interactions are evolutionarily conserved across the entire tree of life. *Nature* **465**, 918–921 (2010).
21. J. A. Dunne, R. J. Williams, N. D. Martinez, R. A. Wood, D. H. Erwin, Compilation and network analyses of Cambrian food webs. *PLoS Biol.* **6**, 693–708 (2008).
22. J. C. S. Nascimento, F. Blanco, M. S. Domingo, J. L. Cantalapiedra, M. M. Pires, The reorganization of predator–prey networks over 20 million years explains extinction patterns of mammalian carnivores. *Ecol. Lett.* **27**, e14448 (2024).
23. M. J. Landis, N. J. Matzke, B. R. Moore, J. P. Huelsenbeck, Bayesian analysis of biogeography when the number of areas is large. *Syst. Biol.* **62**, 789–804 (2013).
24. M. P. Braga, M. J. Landis, S. Nylin, N. Janz, F. Ronquist, Bayesian inference of ancestral host-parasite interactions under a phylogenetic model of host repertoire evolution. *Syst. Biol.* **69**, 1149–1162 (2020).
25. M. P. Braga, N. Janz, S. Nylin, F. Ronquist, M. J. Landis, Phylogenetic reconstruction of ancestral ecological networks through time for pierid

butterflies and their host plants. *Ecol. Lett.* **24**, 2134–2145 (2021).
26. G. V. Dalla Riva, D. B. Stouffer, Exploring the evolutionary signature of food webs' backbones using functional traits. *Oikos* **125**, 446–456 (2016).
27. T. Strydom, *et al.*, Food web reconstruction through phylogenetic transfer of low-rank network representation. *Methods Ecol. Evol.* **2022**, 1–12 (2022).
28. M.-A. Selosse, F. Le Tacon, The land flora: a phototroph-fungus partnership? *Trends Ecol. Evol.* **13**, 15–20 (1998).
29. Y. Zeng, J. J. Wiens, Do mutualistic interactions last longer than antagonistic interactions? *Proc. R. Soc. B Biol. Sci.* **288** (2021).
30. D. Peris, *et al.*, Evolutionary implications of a deep-time perspective on insect pollination. *Biol. Rev.* **100**, 1452–1466 (2025).
31. S. L. Nuismer, L. J. Harmon, Predicting rates of interspecific interaction from phylogenetic trees. *Ecol. Lett.* **18**, 17–27 (2015).
32. F. Massol, E. Macke, M. Callens, E. Decaestecker, A methodological framework to analyse determinants of host–microbiota networks, with an application to the relationships between Daphnia magna's gut microbiota and bacterioplankton. *J. Anim. Ecol.* **90**, 102–119 (2021).
33. W. R. Rimington, S. Pressel, J. G. Duckett, K. J. Field, M. I. Bidartondo, Evolution and networks in ancient and widespread symbioses between Mucoromycotina and liverworts. *Mycorrhiza* **29**, 551–565 (2019).
34. B. Perez-Lamarque, *et al.*, Analysing diversification dynamics using barcoding data: The case of an obligate mycorrhizal symbiont. *Mol. Ecol.* **31**, 3496–3512 (2022).
35. F. Lutzoni, *et al.*, Contemporaneous radiations of fungi and plants linked to symbiosis. *Nat. Commun.* **9**, 1–11 (2018).
36. B. Pichon, R. Le Goff, H. Morlon, B. Perez-Lamarque, Telling mutualistic and antagonistic ecological networks apart by learning their multiscale structure. *Methods Ecol. Evol.* **15**, 1113–1128 (2024).
37. L. Fuzessy, M. A. Pizo, NeoFrugivory: A comprehensive database for frugivory interactions and functional traits in the Neotropics. *Ecology* **106**, e70049 (2025).
38. E. L. Rezende, P. Jordano, J. Bascompte, Effects of phenotypic complementarity and phylogeny on the nested structure of mutualistic networks. *Oikos* **116**, 1919–1929 (2007).
39. B. Perez-Lamarque, H. Morlon, Distinguishing cophylogenetic signal from phylogenetic congruence clarifies the interplay between evolutionary history and species interactions. *Syst. Biol.* **73**, 613–622 (2024).
40. F. Encinas-Viso, T. A. Revilla, R. S. Etienne, Phenology drives mutualistic

network structure and diversity. *Ecol. Lett.* **15**, 198–208 (2012).
41. L. J. Revell, L. J. Harmon, D. C. Collar, Phylogenetic signal, evolutionary process, and rate. *Syst. Biol.* **57**, 591–601 (2008).
42. J. P. Drury, *et al.*, Contrasting impacts of competition on ecological and social trait evolution in songbirds. *PLoS Biol.* **16**, e2003563 (2018).
43. C. Llopis-Belenguer, *et al.*, Sensitivity of bipartite network analyses to incomplete sampling and taxonomic uncertainty. *Ecology* **104**, e3974 (2023).
44. O. Maliet, H. Morlon, Fast and accurate estimation of species-specific diversification rates using data augmentation. *Syst. Biol.* **71**, 353–366 (2022).
45. I. Quintero, N. Lartillot, H. Morlon, Imbalanced speciation pulses sustain the radiation of mammals. *Science (80-.).* **384**, 1007–1012 (2024).
46. I. Junk, *et al.*, Archived natural DNA samplers reveal four decades of biodiversity change across the tree of life. *Nat. Ecol. Evol.*, 1–12 (2025).
47. T. Poisot, *et al.*, Network embedding unveils the hidden interactions in the mammalian virome. *Patterns* **4**, 100738 (2023).
48. J. Clavel, G. Escarguel, G. Merceron, mvMORPH: An R package for fitting multivariate evolutionary models to morphometric data. *Methods Ecol. Evol.* **6**, 1311–1319 (2015).
49. E. P. Martins, T. F. Hansen, Phylogenies and the comparative method: A general approach to incorporating phylogenetic information into the analysis of interspecific data. *Am. Nat.* **149**, 646–667 (1997).
50. J. Clavel, L. Aristide, H. Morlon, A penalized likelihood framework for high-dimensional phylogenetic comparative methods and an application to New-World monkeys brain evolution. *Syst. Biol.* **68**, 93–116 (2019).
51. C. F. Dormann, B. Gruber, J. Fründ, Introducing the bipartite package: analysing ecological networks. *R News* **8**, 8–11 (2008).
52. B. Perez-Lamarque, *et al.*, Do closely related species interact with similar partners? Testing for phylogenetic signal in bipartite interaction networks. *Peer Community J.* **2**, e59 (2022).
53. D. M. de Vienne, *et al.*, Cospeciation vs host-shift speciation: Methods for testing, evidence from natural associations and relation to coevolution. *New Phytol.* **198**, 347–385 (2013).
54. N. E. Langmore, *et al.*, Coevolution with hosts underpins speciation in brood-parasitic cuckoos. *Science (80-.).* **384**, 1030–1036 (2024).
55. E. L. Rezende, J. E. Lavabre, P. R. Guimarães, P. Jordano, J. Bascompte, Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature* **448**, 925–928 (2007).

56. R Core Team, R: A language and environment for statistical computing. (2024).
57. L. L. Sánchez-Reyes, H. Morlon, S. Magallón, Uncovering higher-taxon diversification dynamics from clade age and species-richness data. *Syst. Biol.* **66**, 367–378 (2017).
58. M. Öpik, *et al.*, The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). *New Phytol.* **188**, 223–241 (2010).
59. Y. Jin, H. Qian, V.PhyloMaker2: An updated and enlarged R package that can generate very large phylogenies for vascular plants. *Plant Divers.* **44**, 335–339 (2022).
60. H. Morlon, *et al.*, RPANDA: An R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol.* **7**, 589–597 (2016).
61. K. González-Gutiérrez, J. H. Castaño, J. Pérez-Torres, H. R. Mosquera-Mosquera, Structure and roles in pollination networks between phyllostomid bats and flowers: a systematic review for the Americas. *Mamm. Biol.* **102**, 21–49 (2022).
62. N. S. Upham, J. A. Esselstyn, W. Jetz, Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biol.* **17**, e3000494 (2019).
63. T. Datzmann, O. Von Helversen, F. Mayer, Evolution of nectarivory in phyllostomid bats (Phyllostomidae Gray, 1825, Chiroptera: Mammalia). *BMC Evol. Biol.* **10**, 1–14 (2010).
64. T. H. Fleming, C. Geiselman, W. J. Kress, The evolution of bat pollination: a phylogenetic perspective. *Ann. Bot.* **104**, 1017–43 (2009).
65. W. Jetz, G. H. Thomas, J. B. Joy, K. Hartmann, A. O. Mooers, The global diversity of birds in space and time. *Nature* **491**, 444–448 (2012).
66. W. D. Kissling, K. Böhning-Gaese, W. Jetz, The global distribution of frugivory in birds. *Glob. Ecol. Biogeogr.* **18**, 150–162 (2009).
67. P. Vargas, R. Heleno, J. M. Costa, EuDiS - A comprehensive database of the seed dispersal syndromes of the European flora. *Biodivers. data J.* **11**, e104079 (2023).
68. O. Maliet, F. Hartig, H. Morlon, A model with many small shifts for estimating species-specific diversification rates. *Nat. Ecol. Evol.* **3**, 1086–1092 (2019).