LETTER

Indirect effects drive coevolution in mutualistic networks

Paulo R. Guimarães Jr[1], Mathias M. Pires[2], Pedro Jordano[3], Jordi Bascompte[4] & John N. Thompson[5]

Ecological interactions have been acknowledged to play a key role in shaping biodiversity[1,2]. Yet a major challenge for evolutionary biology is to understand the role of ecological interactions in shaping trait evolution when progressing from pairs of interacting species to multispecies interaction networks[2]. Here we introduce an approach that integrates coevolutionary dynamics and network structure. Our results show that non-interacting species can be as important as directly interacting species in shaping coevolution within mutualistic assemblages. The contribution of indirect effects differs among types of mutualism. Indirect effects are more likely to predominate in nested, species-rich networks formed by multiple-partner mutualisms, such as pollination or seed dispersal by animals, than in small and modular networks formed by intimate mutualisms, such as those between host plants and their protective ants. Coevolutionary pathways of indirect effects favour ongoing trait evolution by promoting slow but continuous reorganization of the adaptive landscape of mutualistic partners under changing environments. Our results show that coevolution can be a major process shaping species traits throughout ecological networks. These findings expand our understanding of how evolution driven by interactions occurs through the interplay of selection pressures moving along multiple direct and indirect pathways.

Coevolution, the reciprocal adaptation resulting from ecological interactions, shapes the adaptive peaks of pairs of interacting species (Fig. 1a, b). Ultimately, selection driven by ecological interactions fuels adaptation in populations[3], affects ecosystems[4], and shapes the responses of ecological assemblages to environmental change[2]. An important challenge in advancing our understanding of how ecological interactions shape biodiversity, however, is to determine how coevolution acts when progressing from pairs or small groups of interacting species[5–12] to species-rich networks[2,13–16]. In species-rich networks, the effects of selection may cascade and produce indirect effects (evolutionary changes prompted by species that are not linked directly as interacting partners). For example, selection imposed by one pollinator species may promote evolutionary changes in a plant species, which may lead to changes in another pollinator species. Indirect effects may change the adaptive landscape (Fig. 1c) and thereby drive trait distributions in biological communities[14–17]. Nevertheless, the lack of a conceptual framework to explore how direct and indirect effects shape trait evolution in species-rich networks has hindered progress in our understanding of these combined effects.

Here, we introduce an approach that combines evolutionary theory and network theory to evaluate how direct and indirect effects shape trait evolution in mutualistic networks. Using a single-trait coevolutionary model, we explored how the importance of indirect effects to trait evolution is affected by selection driven by mutualistic partners and the environment, the additive genetic variance, the slope of the selection gradient, and the constraints imposed by other traits that limit the occurrence of certain interactions (Methods).

We used the structures of 75 empirical networks to parameterize our simulations (Methods). The dataset encompassed six types of mutualism categorized into two broad classes. Intimate mutualisms, such as myrmecophytes hosting ants and anemones hosting anemonefishes, are interactions in which an organism completes at least one life stage on a single host. By contrast, multiple-partner mutualisms are those in which an individual interacts with multiple partners throughout its life, such as pollination by animals, seed dispersal by vertebrates, cleaning interactions involving fishes and shrimps, and the facultative protection of plants by visiting ants. We focused on intimate and multiple-partner interactions because they have been proposed to have disparate effects on coevolution[2,18].

In mutualistic networks, most species interact with a small subset of available partners, which limits the sources of direct evolutionary effects (Fig. 2a). These direct effects, however, may cascade through the network and generate indirect effects. By combining numerical simulations with an analytical approximation, we computed a matrix (T-matrix) that describes how selection imposed directly and indirectly by mutualisms cascades through multiple pathways, thereby reshaping the adaptive landscape (Fig. 2b and Methods). The T-matrix allowed us to partition the potential contributions of direct and indirect effects to trait evolution in species-rich networks (Methods). Because reciprocal selection is a major component of indirect effects, we called the T-matrix the coevolutionary matrix. Sensitivity analyses under different model assumptions showed that the coevolutionary matrix emerges regardless of multiple modelling choices (Supplementary Methods).

Indirect effects contributed strongly to the trait evolution of species in mutualistic networks. This result occurs because the combined weak indirect effects of multiple non-interacting species compensate for the fewer stronger direct effects of interacting partners (Fig. 2b). Surprisingly, indirect evolutionary effects were stronger in species with fewer interactions (specialists) than in highly connected generalists (Fig. 2c, d). For example, in a simulation using a species-rich seed-dispersal network and assuming strong mutualistic selection ($<m> = 0.7 \pm 0.01$, Methods), less than 30% of selective effects on specialist species (yellow dots; Fig. 2d) were driven by their direct partners, whereas the combined effects of non-interacting species accounted for approximately 40% of the selective effects on the traits of specialists. These simulations varying the level of mutualistic selection indicated that indirect effects could be ignored only if mutualisms were a negligible source of selection (Fig. 3a). Our results suggest that coevolution in mutualistic networks may be governed to an unexpectedly large extent by species that do not interact directly with each other.

Having shown the potential importance of indirect effects for coevolution within mutualistic networks, we next investigated how the contribution of indirect effects to trait evolution may differ among types of mutualism. We used empirical networks of different mutualisms to designate which interactions occurred and which did not, but we

[1]Departamento de Ecologia, Instituto de Biociências, Universidade de São Paulo, Rua do Matão 321, Travessa 14, São Paulo - SP 05508-090, Brazil. [2]Departamento de Biologia Animal, Instituto de Biologia, Universidade Estadual de Campinas, Rua Monteiro Lobato 255, Campinas - SP 13083-862, Brazil. [3]Estación Biológica de Doñana (EBD-CSIC), Avenida Americo Vespucio 26, 41092 Sevilla, Spain. [4]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. [5]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, California 95064, USA.
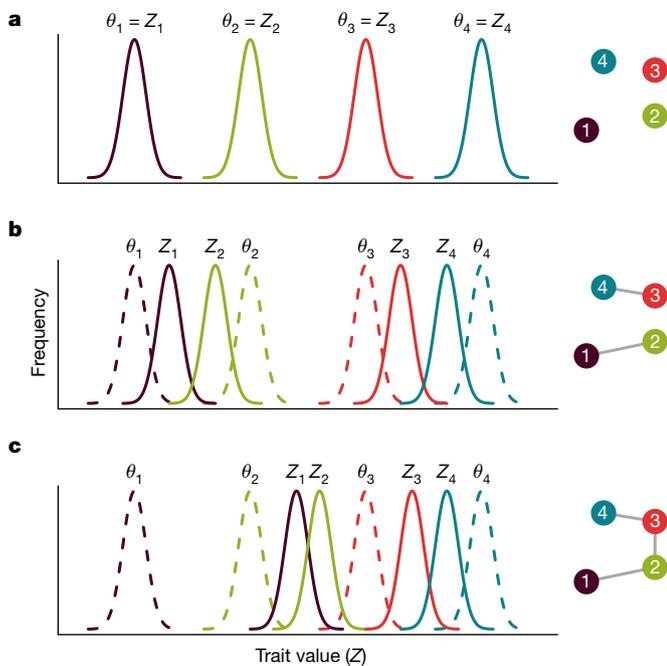
Figure 1 | Mutualisms, selection, and trait distributions. Hypothetical curves describing the distribution of traits for four species in a given community in which the environment imposes selection favouring particular trait values for each species (θ_1, θ_2, θ_3, θ_4). **a**, In the absence of selection imposed by mutualisms, selection favours the mean trait values of each species that match the environmental optima ($Z_1^* = \theta_1$, $Z_2^* = \theta_2$, $Z_3^* = \theta_3$, $Z_4^* = \theta_4$). **b**, Pairwise interactions reshape the adaptive landscape, changing the favoured mean trait values for interacting species. **c**, If species are part of a mutualistic network, then the indirect effects provided by non-interacting species also play a role in reshaping trait distributions.

allowed interaction strengths to evolve in response to trait evolution. These interactions connecting species in the network drive the direct evolutionary effects, but also create pathways connecting non-interacting species. Indirect evolutionary effects varied among different types of mutualism (general linear model, $F_{5,593} = 194.94$, $P < 0.0001$; Supplementary Table 2), exhibiting greater influence in multiple-partner mutualisms than in mutualisms involving more intimate interactions (Fig. 3b). These differences between types of mutualism held after controlling for the effects of species richness (Supplementary Table 3) and when using empirical ecological dependences as estimates of direct evolutionary effects (Supplementary Methods).

We then evaluated which network patterns favour indirect evolutionary effects (Methods). Multiple-partner mutualisms often form nested networks, whereas intimate mutualisms form highly modular networks[18]. Indirect evolutionary effects were weaker in modular networks than in species-rich, nested networks ($R^2 = 0.69$, $F_{1,73} = 160.91$, $P < 0.0001$; Fig. 3c). An analytical approximation of the model using spectral graph theory confirmed that nestedness and species richness increase indirect effects by enriching the number of pathways connecting species in the network (Supplementary Methods).

In a multispecies network, environmental changes may favour evolutionary responses in opposing directions for different species. If indirect effects contribute extensively to trait evolution, then these opposing indirect effects may spread through the network and create conflicting coevolutionary cascades. We therefore evaluated how indirect effects may shape evolutionary responses following environmental perturbations in mutualistic networks. We ran the model until it reached equilibrium and then simulated a long-lasting change in environmental selection, such as a sustained change in rainfall, by displacing the environmental optimum (Methods).

The perturbations changed the adaptive landscape, which triggered changes in the network organization by altering the evolutionary effects
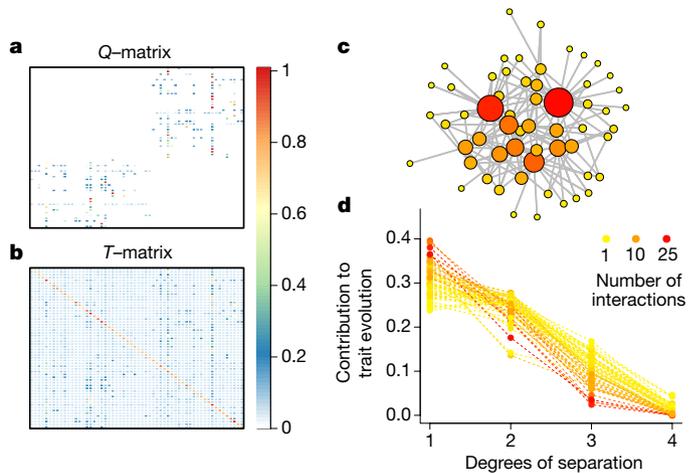

Figure 2 | Direct and indirect effects in networks. a, b, Matrices representing a mutualistic network (network 45; Supplementary Table 1, see also Methods), with warmer colours indicating stronger evolutionary effects of a species (column) on the trait evolution of another species (row). **a**, Mutualistic networks often show few direct evolutionary effects (elements of Q-matrix) but (**b**) numerous indirect effects (elements of T-matrix) because multiple pathways connect species. **c, d**, In species-rich networks, specialists (smaller nodes with cooler colours) are more affected than generalists (larger nodes with warmer colours) by the indirect evolutionary effects of species two or more degrees of separation away. Parameters: $\varphi = 0.2 \pm 0.01$, $\theta_i = U[0, 10]$, and $\langle m \rangle = 0.7 \pm 0.01$.

of interactions and subsequently promoted further trait evolution. The greater the indirect evolutionary effects, the longer was the time required for a given mutualistic network to reach a new equilibrium ($R^2 = 0.71$, $F_{1,73} = 179.68$, $P < 0.001$; Fig. 4a). Thus, indirect effects induced long transients[19] in the coevolutionary dynamics of multiple-partner mutualisms (Fig. 4a). The conflicting indirect effects also reduced the rate of directional evolution ($R^2 = 0.69$, $F_{1,73} = 165.56$, $P < 0.001$, Fig. 4b). These results held after controlling for the effects of species richness (Supplementary Table 4). Thus, the structure of multiple-partner mutualisms may slow the response to environmental changes while lengthening the cascading effects of the perturbations, thereby promoting the endurance of coevolutionary dynamics.

Although indirect effects are a major component of how species interactions affect biodiversity, as in the case of trophic cascades[20], trait-mediated cascades[12], and competition[21], we are just beginning to understand how much indirect effects contribute to evolution in species-rich assemblages. Quantifying indirect effects in complex networks is a current challenge in many fields of research. Indirect effects are a fundamental component of processes affecting population genetic structure[22], financial markets[23], and cultural practices[24]. The framework presented here, combined with appropriate models for each system, may contribute to the characterization of indirect effects in other complex networks. By introducing this framework, we have shown that indirect evolutionary effects[12] are a pervasive consequence of nestedness and species richness in multiple-partner mutualisms.

Our results on the role of indirect effects in shaping trait evolution suggest three major properties of coevolution in mutualistic networks. First, these results challenge the view that only intimate mutualisms are highly coevolved. Rather, coevolution proceeds differently in different types of interaction. In intimate mutualisms, the direct selection imposed by the mutual interdependence between partners is expected to be strong[2,18], whereas the organization of these small networks constrains the strength of indirect effects. By contrast, our analysis suggests that indirect effects markedly affect trait evolution in multiple-partner mutualisms. Thus, selection regimes imposed by multiple-partner interactions are the outcome of a complex interplay among conflicting selection pressures operating through multiple pathways. This high
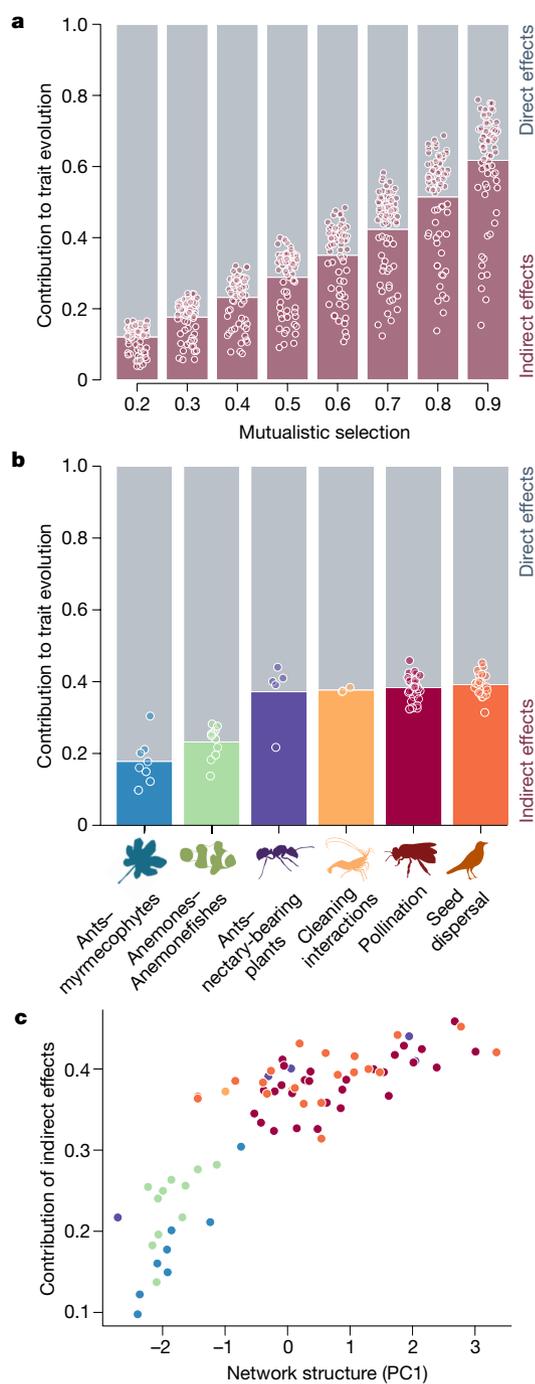
Figure 3 | Determinants of indirect effects. a–c, Each circle represents the mean contribution of indirect effects to trait evolution ($n = 800$ simulations) parameterized with a given empirical network ($n = 75$ empirical networks). **a,** The mean contribution of indirect effects (red bars) increases with mutualistic selection. **b,** After controlling for the level of mutualistic selection, the indirect effects varied across mutualisms. **c,** Indirect effects increased along a gradient based on PC1 of a principal component analysis (Methods) from small, modular networks (negative values) to species-rich, nested networks (positive values). Types of mutualism (**b, c**): blue, ants–myrmecophytes; green, anemones–anemonefishes; purple, ants–nectary-bearing plants; light orange, cleaning interactions; red, pollination; dark orange, seed dispersal.

level of integration may provide a mechanism for the emergence of community-level trait patterns in mega-diversified mutualisms[17].

Second, indirect effects may alter the consequences of environmental selective pressures by triggering trait evolution in other species
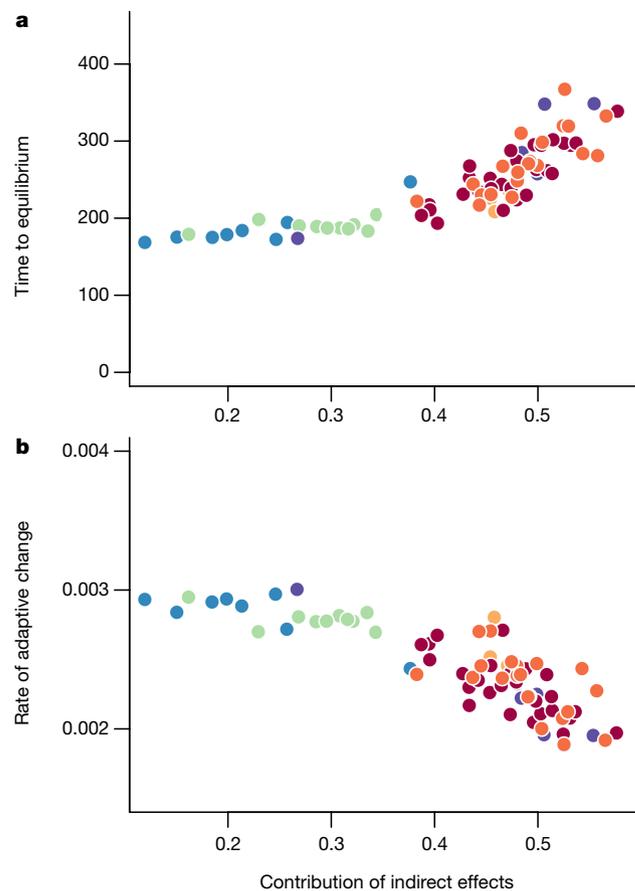
Figure 4 | Indirect effects and environmental change. a, Higher contributions of indirect effects to trait evolution (least square estimates of a general linear model; see Methods) led to longer durations of the coevolutionary dynamics triggered by the environmental changes. **b,** Longer coevolutionary dynamics resulted from the contribution of indirect effects that slowed the rate of directional adaptive change in the mean trait values of interacting species, measured as the mean amount of trait change per species per time step. Different colours indicate distinct types of mutualism (Fig. 3).

within a mutualistic network. In multiple-partner mutualisms, in which conflicting indirect effects are likely to be pervasive, the effects of environmental change would take longer to diminish than they would in intimate mutualisms. Because biological communities are continuously affected by perturbations, these results imply that the traits of species immersed in species-rich mutualistic networks may often be far from equilibrium[19], leading to slow but continuous coevolution that will repeatedly reshape selection regimes and species traits at a local level.

Third, because interactions are mediated by species traits, trait evolution leads to the reorganization of the network, generating eco-evolutionary feedbacks. The reorganization of the network structure as a result of the indirect effects of coevolution may explain why and how mutualisms persist amid the turnover of species and interactions across space and time[25]. Conflicting indirect effects, though, may lessen the degree and slow the rate at which species respond to the rapid environmental change currently driven by human activities[26], affecting population demography in ways that increase their vulnerability to extinction[3,27]. If this is true, then the network properties that favour direct and indirect evolutionary effects, continuously reshaping species traits under normal conditions, may also threaten these same species when the community is subjected to the rapid, human-driven environmental changes occurring in biological communities worldwide.

1. Ehrlich, P. R. & Raven, P. H. Butterflies and plants: a study in coevolution. *Evolution* **18,** 586–608 (1964).
2. Thompson, J. N. *The Geographic Mosaic of Coevolution* (Univ. Chicago Press, 2005).
3. Galetti, M. *et al.* Functional extinction of birds drives rapid evolutionary changes in seed size. *Science* **340,** 1086–1090 (2013).
4. Koskella, B. & Brockhurst, M. A. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol. Rev.* **38,** 916–931 (2014).
5. Parchman, T. L. & Benkman, C. W. Diversifying coevolution between crossbills and black spruce on Newfoundland. *Evolution* **56,** 1663–1672 (2002).
6. Brodie, E. D. III. Genetic correlations between morphology and antipredator behaviour in natural populations of the garter snake *Thamnophis ordinoides*. *Nature* **342,** 542–543 (1989).
7. Ridenhour, B. J. Identification of selective sources: partitioning selection based on interactions. *Am. Nat.* **166,** 12–25 (2005).
8. Strauss, S. Y., Sahli, H. & Conner, J. K. Toward a more trait-centered approach to diffuse (co)evolution. *New Phytol.* **165,** 81–90 (2005).
9. Iwao, K. & Rausher, M. D. Evolution of plant resistance to multiple herbivores: quantifying diffuse coevolution. *Am. Nat.* **149,** 316–335 (1997).
10. Thompson, J. N., Schwind, C., Guimarães, P. R. Jr & Friberg, M. Diversification through multitrait evolution in a coevolving interaction. *Proc. Natl Acad. Sci. USA* **110,** 11487–11492 (2013).
11. Fox, L. R. Diffuse coevolution within complex communities. *Ecology* **69,** 906–907 (1988).
12. Ohgushi, T., Schmitz, O. & Holt, R. D. *Trait-Mediated Indirect Interactions: Ecological and Evolutionary Perspectives* (Cambridge Univ. Press, 2012).
13. Gómez, J. M., Perfectti, F., Bosch, J. & Camacho, J. P. M. A geographic selection mosaic in a generalized plant-pollinator-herbivore system. *Ecol. Monogr.* **79,** 245–263 (2009).
14. Nuismer, S. L., Jordano, P. & Bascompte, J. Coevolution and the architecture of mutualistic networks. *Evolution* **67,** 338–354 (2013).
15. Santamaría, L. & Rodríguez-Gironés, M. A. Linkage rules for plant-pollinator networks: trait complementarity or exploitation barriers? *PLoS Biol.* **5,** e31 (2007).
16. Guimarães, P. R. Jr, Jordano, P. & Thompson, J. N. Evolution and coevolution in mutualistic networks. *Ecol. Lett.* **14,** 877–885 (2011).
17. Elias, M., Gompert, Z., Jiggins, C. & Willmott, K. Mutualistic interactions drive ecological niche convergence in a diverse butterfly community. *PLoS Biol.* **6,** 2642–2649 (2008).
18. Fontaine, C. *et al.* The ecological and evolutionary implications of merging different types of networks. *Ecol. Lett.* **14,** 1170–1181 (2011).
19. Hastings, A. Transients: the key to long-term ecological understanding? *Trends Ecol. Evol.* **19,** 39–45 (2004).
20. Terborgh, T. & Estes, J. A. *Trophic Cascades: Predators, Prey, and the Changing Dynamics of Nature* (Island, 2010).
21. Levine, J. M., Bascompte, J., Adler, P. B. & Allesina, S. Beyond pairwise mechanisms of species coexistence in complex communities. *Nature* **546,** 56–64 (2017).
22. Dyer, R. J. & Nason, J. D. Population graphs: the graph theoretic shape of genetic structure. *Mol. Ecol.* **13,** 1713–1727 (2004).
23. Haldane, A. G. & May, R. M. Systemic risk in banking ecosystems. *Nature* **469,** 351–355 (2011).
24. Fowler, J. H. & Christakis, N. A. Cooperative behavior cascades in human social networks. *Proc. Natl Acad. Sci. USA* **107,** 5334–5338 (2010).
25. Eriksson, O. Evolution of angiosperm seed disperser mutualisms: the timing of origins and their consequences for coevolutionary interactions between angiosperms and frugivores. *Biol. Rev. Camb. Philos. Soc.* **91,** 168–186 (2016).
26. Tylianakis, J. M., Didham, R. K., Bascompte, J. & Wardle, D. A. Global change and species interactions in terrestrial ecosystems. *Ecol. Lett.* **11,** 1351–1363 (2008).
27. Miller-Struttmann, N. E. *et al.* Functional mismatch in a bumble bee pollination mutualism under climate change. *Science* **349,** 1541–1544 (2015).

METHODS

No statistical methods were used to predetermine sample size.

Coevolutionary model. We used a coevolutionary model for mutualistic networks to simulate the phenotypic evolution of interacting species. This model incorporates selection gradients into a weighted network framework in which mutualistic interactions vary in interaction strength over time. The incorporation of selection gradients allowed us to connect trait evolution explicitly with the mean fitness consequences of mutualistic interactions. The mean trait evolution of species i is described as follows:

$$Z_i^{(t+1)} = Z_i^{(t)} + \varphi_i \left[\sum_{j,j\neq i}^{N} q_{ij}^{(t)} \left(Z_j^{(t)} - Z_i^{(t)} \right) + \left(1 - \sum_{j,j\neq i}^{N} q_{ij}^{(t)} \right) \left(\theta_i - Z_i^{(t)} \right) \right] \quad (1)$$

where $Z_i^{(t)}$ is the mean trait value of the local population of species i at time t, φ_i is a compound parameter that affects the slope of the selection gradient and is proportional to the additive genetic variance (Supplementary Methods), N is the number of species in the network, $q_{ij}^{(t)}$ is the interaction weight that describes the evolutionary effect of species j on the selection gradient of species i and evolves over time because of trait matching, $1 - \sum_{j,j\neq i}^{N} q_{ij}^{(t)}$ weights the relative importance of selective pressures that are not related to mutualisms (hereafter environmental selection) in shaping the adaptive landscape, and θ_i is the phenotype favoured by the environmental selection (see the Supplementary Methods for the complete derivation of the model). Evolutionary effects, $q_{ij}^{(t)}$, were defined as a function of trait matching:

$$q_{ij}^{(t)} = m_i \frac{a_{ij} e^{-\alpha\left(Z_j^{(t)} - Z_i^{(t)}\right)^2}}{\sum_{k, i\neq k}^{N} a_{ik} e^{-\alpha\left(Z_k^{(t)} - Z_i^{(t)}\right)^2}} \quad (2)$$

where α is a scaling constant that controls the sensitivity of the evolutionary effect to trait matching ($\alpha = 0.2$ in all simulations). The parameter m_i is the level of mutualistic selection, $(m_i = \sum_{j,j\neq i}^{N} q_{ij}^{(t)})$. By varying m_i, we can explore how the strength of mutualistic selection affects the contribution of indirect effects. The parameter a_{ij} is an element of A, which is the adjacency matrix of a given mutualistic network with N species, and $a_{ij} = 1$ if species i and j are a pair of interacting species and 0 otherwise. Because the mutualisms analysed here involve two sets of species, S_a and S_b (for example, pollinators and plants) and because the interactions occur only between species from different sets, A has the following form:

$$A = \begin{pmatrix} 0_{N_a \times N_a} & R_{N_a \times N_b} \\ R_{N_b \times N_a}^T & 0_{N_b \times N_b} \end{pmatrix} \quad (3)$$

where N_a is the number of species of set S_a (for example, pollinators); N_b is the number of species of set S_b (for example, flowering plants), $N = N_a + N_b$, 0 is a submatrix in which all elements are zero, and $R_{N_a \times N_b}$ is the biadjacency matrix that defines a bipartite graph formed by two sets of species. If $a_{ij} = 0$ then $q_{ij}^{(t)} = 0$ for any t, and the interaction between species j and species i is a forbidden link. A forbidden link is one in which two species cannot interact because, for example, traits in one or both species prevent the interaction. We used information on empirical networks in our dataset to parameterize the matrix A, see below. As a first approximation, we assumed all $a_{ij} = 0$ in empirical networks are fixed forbidden links. Later we relaxed this assumption allowing interaction rewiring due to coevolution (Supplementary Methods). The model is available as a MATLAB script upon request.

Dataset. Our dataset included a wide range of terrestrial and marine mutualistic networks that vary considerably in their natural history attributes and network patterns (see Supplementary Table 1, and the references therein). Interactions between anemones and anemonefishes and between myrmecophytes and their sheltered ant colonies are examples of intimate mutualisms in which individuals create sustained interactions and form small, highly modular networks[18]. By contrast, the cleaning interactions in coral reefs, seed dispersal by vertebrates, pollination by animals, and interactions between nectary-bearing plants and their protective ants are representatives of multiple-partner mutualisms among free-living species in which individuals may interact with dozens or hundreds of different partners across a lifetime[2,18] and often form species-rich, nested networks. For a subset of mutualistic networks (38 networks), we also obtained information on the frequency of the interactions. We describe the structural patterns observed in these networks in Supplementary Table 1.

Model parameterization and numerical simulations. We used empirical information on the 75 mutualistic networks to parameterize the matrix A in numerical simulations such that the absence of any observed interactions between two species was assumed to represent an evolutionary forbidden link, $a_{ij} = 0$. In empirical networks, the absence of interactions may represent forbidden links or interactions that were not observed owing to the inherent challenges of sampling ecological interactions. However, we are not interested in each network but rather in major differences in the contribution of indirect effects to coevolution across different network patterns. Other model parameters were sampled from statistical distributions: $Z_i^{(0)}$ and θ_i were sampled from uniform distributions with the range [0, 10], φ_i was sampled from a truncated normal distribution with a mean $\pm$ s.d. $= 0.2 \pm 0.01$ and bounded between 0 and 1, and the level of mutualistic selection for a given species, m_i, was sampled from truncated normal distributions with a mean $\pm$ s.d. $= <m> \pm 0.01$ and bounded between 0 and 1, where $<m>$ is the mean level of mutualistic selection in the network. We explored how the mean level of mutualistic selection affected the dynamics by exploring $<m>$ ranging from 0.2 to 0.9 (in increments of 0.1). We performed 100 simulations per combination of $<m>$ (eight different values) and empirical network (75 mutualistic networks), for a total of 6×10^4 simulations. Each simulation ended after the species achieved asymptotic trait values, which were defined as mean changes in species trait values at sequential time steps $<10^{-6}$.

Matrix of direct evolutionary effects (Q-matrix). Networks can be represented as adjacency matrices in which row and column i represent species i and non-empty matrix elements depict pairwise interactions (Fig. 2a). In our model, $q_{ij}^{(t)}$ describes the direct evolutionary effects of selection imposed by species j in trait evolution of species i (see the Coevolutionary model section in the Methods). Consequently, Q describes all the direct evolutionary effects of mutualistic interactions in a network. In all simulations, evolutionary dynamics reached an equilibrium at which the trait values and evolutionary effects of species on their partners became fixed (Extended Data Fig. 1), leading to a stationary Q-matrix.

Coevolutionary matrix (T-matrix). Our analytical study showed that the equilibrium of our coevolutionary model is stable and defined as $Z^* = T\Theta$, in which Z^* is an $N \times 1$ vector with the species mean traits at equilibrium, Θ is an $N \times 1$ vector describing the trait values favoured by environmental selection, and the coevolutionary matrix T (Fig. 2b) connects the values favoured by environmental selection (Θ) to the fixed trait values (Z^*) by means of selection imposed by mutualisms. The elements of this matrix indicate the relative contribution of each interacting and non-interacting species in the network (columns) to the selection gradient shaping trait evolution of a given species (row) in the network (Methods). The main diagonal of T describes the effects of environmental selection on trait evolution of a given species. Our analytical study showed that the matrix T is defined as $T = (I - Q)^{-1}\Psi$ for any mutualistic network, in which I is the identity matrix, Q is an $N \times N$ matrix containing the pairwise, direct evolutionary effects of mutualistic interactions (equation (2)), and Ψ is an $N \times N$ diagonal matrix in which $\Psi_{ii} = 1 - m_i$ for any species i. The component $(I - Q)^{-1}$ is the result of the multiple pathways connecting species in the network (Supplementary Methods), which allowed us to estimate the contribution of indirect effects to trait evolution in our simulations. We performed a sensitivity analysis to determine the dependency of T on the parameter values and the model assumptions, including the way we modelled trait–fitness relationships, selection imposed by the environment, evolutionary effects, and forbidden links. The results were robust regardless of the main assumptions of our model. Derivation of the analytical results and details about the sensitivity analysis are available in the Supplementary Methods.

Quantification of indirect effects. We related the elements of matrix T to the adjacency matrix A to estimate the contribution of indirect effects to trait evolution in our simulations. We defined indirect effects as the proportional effect of non-interacting species on the trait evolution of a given species. The relative contribution of indirect effects to trait evolution in a network was defined as $\kappa = \sum_{i}^{N}\sum_{j,i\neq j}^{N}(1 - a_{ij})t_{ij} / \sum_{i}^{N}\sum_{j,i\neq j}^{N}t_{ij}$ in which t_{ij} is an element of matrix T. The relative contribution of direct effects was defined as $\varsigma = 1 - \kappa$.

Indirect effects and different types of mutualism. We compared the relevance of indirect effects to different mutualisms using simulations parameterized with the empirical network structure of different mutualisms. We computed the relative contribution of indirect effects to trait evolution, κ, in each simulation (100 simulations per network and mean level of mutualistic selection, $<m>$). We used the mean κ for each combination of empirical network and $<m>$ as the response variable and we fitted a general linear model using $<m>$ and the type of mutualism as explanatory variables. We then computed the least squares estimates of the mean relative contribution of indirect effects for each type of mutualism, controlling for the level of mutualistic selection, as a measure of the contribution of indirect effects to evolutionary dynamics in the numerical simulations (Supplementary Table 2).

Indirect effects and network structure. We characterized the network structure using four network descriptors. Species richness is the total number of species,

$(N_a + N_b)$. Connectance is the proportion of possible interactions that are actually recorded and is represented by $C = L/N_aN_b$, in which L is the total number of interactions. Modularity is the level to which the network is organized into semi-independent groups of interacting species. Modularity was quantified using the bipartite modularity index Q_E, which was estimated using the program MODULAR under a simulated annealing optimization algorithm[28]. Q_E ranges between 0 (no modularity) and 1 (high modularity) (see ref. 28 and references therein). Nestedness occurs when species in the network interact with a proper subset of the partners of more connected species. Nestedness was quantified using NODF[29]. NODF ranges from 0 (no nestedness) to 100 (perfect nestedness). We controlled for the confounding effects of species richness, connectance, and heterogeneity in the number of interactions per species on modularity and nestedness using a null model approach[30]:

$$Q_S = \frac{Q_E - Q_N}{Q_N} \qquad (4a)$$

$$N_S = \frac{1}{100} \frac{N_E - N_N}{N_N} \qquad (4b)$$

where $Q_S(N_S)$ is the standardized estimate for modularity (nestedness) for an empirical network, $Q_E(N_E)$ is the raw estimate of modularity (nestedness) for an empirical network, and $Q_N(N_N)$ is the mean value of modularity (nestedness) for null model networks ($n = 100$ null model networks). We used a null model in which the probability of species i of set S_a interacting with a given species j of set S_b was $P_{ij} = 0.5[(k_i/N_b) + (k_j/N_a)]$, where $k_i(k_j)$ is the number of interactions of species $i(j)$[30].

We then explored the relationship between indirect effects and network structure. We used a general linear model in which the explanatory variables were network identity and the mean level of mutualistic selection $<m>$, which ranged from 0.2 to 0.9 (in increments of 0.1). The response variable was the relative contribution of indirect effects to trait evolution. We used this model to obtain least squares estimates of the contributions of indirect effects to trait evolution in each network.

Metrics that describe network structure often show strong correlations in their values. We performed a principal component analysis with the four structural descriptors of the 75 mutualistic networks (species richness, connectance, relative nestedness, and relative modularity) to combine the structural information provided by the network metrics into a single descriptor. The first principal component (PC1) contained 56.4% of the variation in the network descriptors and was negatively associated with the relative level of modularity (-0.429) and connectance (-0.398), and positively associated with the relative level of nestedness (0.586) and species richness (0.560). We then used simple linear regressions to test whether the network structure (scores of PC1) affected the indirect effects (the least squares estimates of the contribution of indirect effects in each network). We then explored the effect of each network descriptor on the contribution of indirect effects to trait evolution (Supplementary Table 5).

Environmental change simulation. We ran 100 simulations for each empirical network (parameters: 75 empirical networks; $\varphi = 0.2 \pm 0.01$, $\theta_i = U[0, 10]$, and $<m> = 0.7 \pm 0.01$). For each simulation we first allowed the system to achieve equilibrium, computing the contribution of indirect effects using the T-matrix. We then performed perturbations that led to sustained changes in the selection imposed by the environment (press perturbations). Press perturbations were simulated by changing the phenotype favoured by the environment by an amount ε_i from the θ_i of each species of the network in each simulation. We sampled ε_i from a uniform distribution [0, 1]. We then ran the simulations until a new equilibrium was reached. We tested whether the contribution of non-interacting species to the evolutionary dynamics was associated with the mean time to the new equilibrium and the rate of directional change. The rate of directional change was computed as the mean change in the trait values per species per time step after the perturbation. We also performed an additional set of simulations in which we sampled ε_i from a uniform distribution $[-1, 1]$, which produced similar results (Supplementary Methods).

Code availability. All MATLAB and R scripts used in this study are available from the corresponding author upon request.

Data availability. The dataset used in this paper is available from the corresponding author upon request.

28. Marquitti, F. M. D., Guimarães, P. R., Pires, M. M. & Bittencourt, L. F. MODULAR: software for the autonomous computation of modularity in large network sets. *Ecography* **37,** 221–224 (2014).
29. Almeida-Neto, M., Guimarães, P., Guimarães, P. R., Loyola, R. D. & Ulrich, W. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos* **117,** 1227–1239 (2008).
30. Bascompte, J., Jordano, P., Melián, C. J. & Olesen, J. M. The nested assembly of plant–animal mutualistic networks. *Proc. Natl Acad. Sci. USA* **100,** 9383–9387 (2003).

(a)

(b)
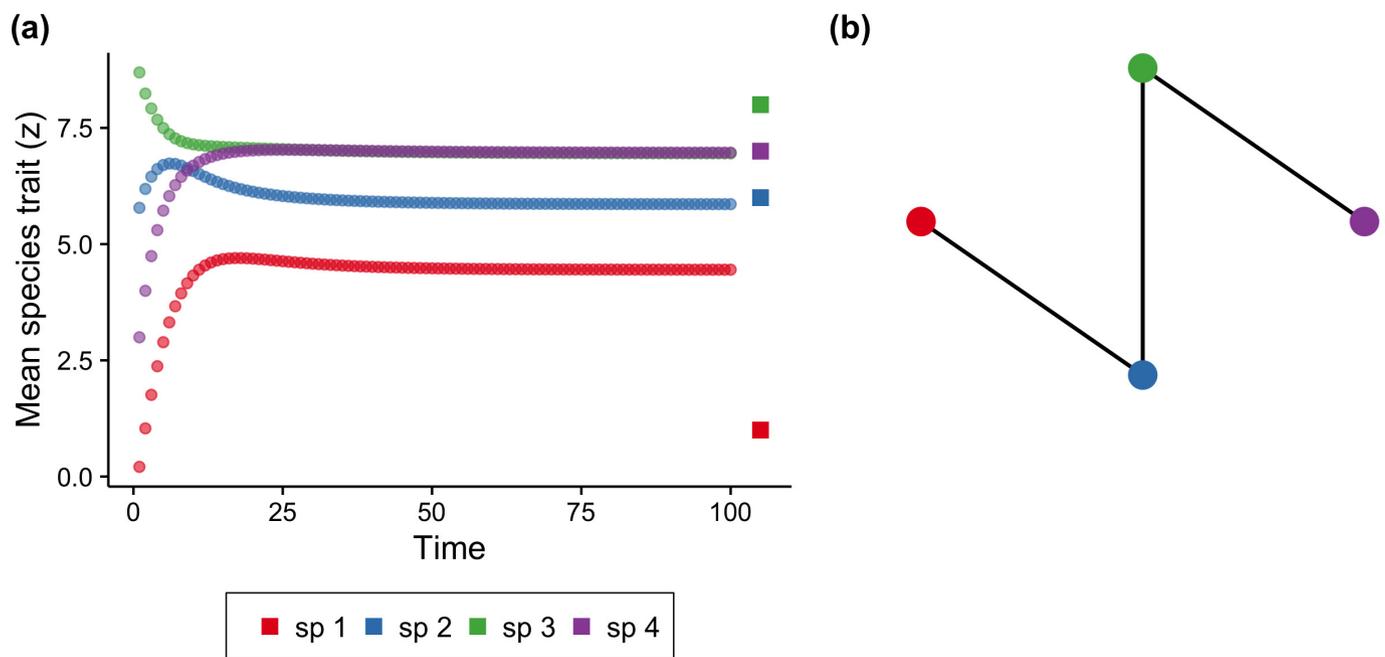
Extended Data Figure 1 | Trait dynamics of a mutualistic network. a, A typical simulation of the coevolutionary model describing the temporal variation in the trait dynamics for a four-species network (**b**, see also Fig. 1c). Points of a given colour represent the evolution of the mean trait value of one species. Small squares indicate the environmental optima of the species in the network. Squares and points corresponding to the same species are presented in the same colour. The mean mutualistic selection was set at $<m> = 0.7 \pm 0.01$. Other parameters: $\varphi = 0.2 \pm 0.01$, $\theta_i = U[0, 10]$. Similarly, the simulations converged to equilibrium for all empirical networks.

naturE research

Corresponding author(s): Guimarães, P.R.

☐ Initial submission ☒ Revised version ☐ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

▶ Experimental design

1. **Sample size**

   Describe how sample size was determined.

   > We computed 100 simulations per combination of empirical network used to parameterize the dataset and level of mutualistic selection - a key parameter of the model. We used 100 simulations because this number allow to characterize the mean values of response variables by minimizing statistical fluctuations due to smaller sample sizes.

2. **Data exclusions**

   Describe any data exclusions.

   > There is no data exclusion.

3. **Replication**

   Describe whether the experimental findings were reliably reproduced.

   > All MATLAB codes and empirical datasets used to parameterize the models are available, allowing reliable reproduction of all of our results.

4. **Randomization**

   Describe how samples/organisms/participants were allocated into experimental groups.

   > Not Applicable to our study.

5. **Blinding**

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   > Not Applicable to our study.

   Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. **Statistical parameters**

   For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The <u>exact sample size</u> (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. P values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A clear description of statistics including <u>central tendency</u> (e.g. median, mean) and <u>variation</u> (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

▶ Software

Policy information about availability of computer code

7. Software

| Describe the software used to analyze the data in this study. | We used manly MATLAB and R scripts we wrote for simulations. We used some well-known, published programs to compute two network statistics. |

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

▶ Materials and reagents

Policy information about availability of materials

8. Materials availability

| Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company. | All datasets and codes are available. |

9. Antibodies

| Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species). | Not Applicable. |

10. Eukaryotic cell lines

| a. State the source of each eukaryotic cell line used. | Not Applicable. |

| b. Describe the method of cell line authentication used. | Not Applicable. |

| c. Report whether the cell lines were tested for mycoplasma contamination. | Not Applicable. |

| d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use. | Not Applicable. |

▶ Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

| Provide details on animals and/or animal-derived materials used in the study. | Not Applicable. |

Policy information about studies involving human research participants

12. Description of human research participants

| Describe the covariate-relevant population characteristics of the human research participants. | Not Applicable. |