JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12908

Research Article

Selection to attract pollinators and to confuse antagonists specializes fig–pollinator chemical communications

Yang Yang[1], Yuan-Ye Zhang[2], Yu Zhang[1], Shan Chen[1], Qian-Ya Li[1], Rong Wang[1,3,4]*, and Xiao-Yong Chen[1,3,4]*

[1]Zhejiang Tiantong Forest Ecosystem National Observation and Research Station, Shanghai Key Laboratory for Urban Ecological Processes and Eco-Restoration, School of Ecological and Environmental Sciences, East China Normal University, Shanghai 200241, China
[2]Key Laboratory of the Ministry of Education for Coastal and Wetland Ecosystems, College of the Environment and Ecology, Xiamen University, Xiamen 361102, Fujian, China
[3]Shanghai Institute of Pollution Control and Ecological Security, Shanghai 200092, China
[4]Shanghai Engineering Research Center of Sustainable Plant Innovation, Shanghai 200231, China
*Authors for correspondence. Rong Wang. E-mail: rwang@des.ecnu.edu.cn; Xiao-Yong Chen. E-mail: xychen@des.ecnu.edu.cn

Received 27 May 2022; Accepted 3 August 2022; Article first published online 5 August 2022

Abstract Chemical communication is critical in establishing angiosperm–pollinator mutualisms. However, our understanding of how chemical communication shapes coevolution remains limited. Here, we integrated information theory to model three coevolutionary scenarios (I–III), where the pollinator fitness is always optimized by the highest certainty of chemical information provided by plants, but plant fitness is determined by (I) the certainty of chemical information attracting pollinators, (II) the uncertainty of chemical information confusing antagonists, or (III) both aspects. We found that the statistical properties of empirical plant volatiles from 45 pairs of fig–pollinator mutualisms were best explained by the selection from both pollinators and antagonists (scenario III). Under this scenario, plant–pollinator mutualisms evolve to be specialized and as few as two volatile chemicals could supply sufficient information for pollinators' host identification. Our study provides new insights into plant–pollinator coevolution and will facilitate further studies on the evolution and diversification in specialized plant–pollinator–herbivore systems.

Key words: coevolution, conditional entropy, fig–pollinator mutualism, fitness, information theory, mutual information, plant–insect interaction.

1 Introduction

Plant–pollinator mutualisms facilitate the coadaptation and codiversification of angiosperms and insects and contribute crucially to the maintenance of biodiversity (Steffan-Dewenter & Westphal, 2008; Johnson, 2010; Vasiliev & Greenwood, 2020). These mutualisms depend critically on chemical communication (Raguso, 2008; Raguso et al., 2015), and chemical signals are mediated through volatile organic compounds (VOCs) of plants. However, existing studies have shown that flowers often emit VOCs that are far more complex than required for pollinators' host identification (Zito et al., 2015; Pichersky & Raguso, 2018; Zhang et al., 2020; Wang et al., 2021). This is at least partly because, in addition to pollination mutualisms, plants are attracted and attacked by a large variety of herbivores (i.e., antagonists), and thus VOC profiles of flowers are under selection to defend from antagonists by increasing complexity (McCall & Irwin, 2006; Raguso, 2016; Zu et al., 2020). Therefore, a comprehensive picture of plant–pollinator chemical communications can help to elucidate how mutualists achieve reciprocal adaptation and coevolution.

Plants emit VOCs to attract pollinators, and when VOCs deliver reliable information (specific VOCs that can induce electrophysiological activities in the pollinators) the plants are easily identified by pollinators (Kessler & Halitschke, 2009). Producing VOCs with reliable information is thus selectively advantageous for plants, and the selective force to attract pollinators will drive the plants to produce these attractive VOCs (Falara et al., 2013; Lukas et al., 2019). Nevertheless, such reliable information is also easily accessible and utilized by antagonists, and an alternative evolutionary force has been raised, where plant species tend to generate similar VOC profiles for preventing antagonists from locating their preferable hosts (i.e., increasing the complexity and uncertainty of information to identify hosts) (Johnson et al., 2015; Schiestl, 2015; Zu et al., 2020, 2022). Therefore, the fitness (the ability of survival and breeding for living organisms) of plants ought to be simultaneously affected by the selection to attract pollinators and to confuse antagonists. Such conflicting selection could lead to an equilibrium in coevolution, and one possible outcome is to form specialized plant–pollinator mutualisms, as the optimized fitness of plants can be reached when plants only change the minimum number of VOCs to ensure pollinators' identification and maximally avoid antagonists' utilization.

However, this hypothesis has never been tested, and a major challenge is to identify the specific attractants from plant VOC profiles, which often involves laborious efforts to separate VOC components and numerous behavioral validation experiments. This has led to the current situation in which important plant VOCs in plant–pollinator communication are known in only a few species (Chen et al., 2009; Proffit et al., 2020; Zhang et al., 2020; Wang et al., 2021). Recently, the application of information theory has provided new insights into chemical communications between plants and insects (Doyle, 2009; Zu et al., 2020). This novel approach uses conditional entropy (higher conditional entropy reflects higher uncertainty and unreliable information) to evaluate whether sufficient chemical information is delivered from plants to target insects, without the prerequisite of knowing the role of each VOC (Zu et al., 2020), and thus can be used to explore the coevolution of specialized plant–pollinator mutualisms. High conditional entropy of plant information indicates a high uncertainty for host identification by pollinators/antagonists, which affects pollination success of plants and the accessibility of nectar for pollinators, thus relating to the survival and breeding (i.e., fitness) of both plants and pollinators.

Integrating the information theory, we tested the role of selection to attract pollinators and to confuse antagonists in the coevolution of specialized plant–pollinator mutualisms using simulations. In the simulations, we did not consider any empirical VOC data in insect–plant interactions and randomly generated mutations of plants' VOCs, which are selected under three coevolutionary scenarios (I–III) (Fig. 1). In scenario I, we assume plant fitness is selected by plants' ability to attract pollinators, and chemical information with high certainty (low entropy) is advantageous. The plant fitness is accordingly defined as reversely related to conditional entropy of plants to VOCs ($1 - H(P|V)$). In scenario II, plant fitness is selected by the plants' ability to confuse antagonists, and thus plants benefit from high uncertainty of chemical information with their fitness being represented by the conditional entropy of plants to VOCs ($H(P|V)$). In scenario III, we expect that plant fitness is selected by both the plants' ability to attract pollinators and to confuse antagonists, leading to a trade-off in the uncertainty of chemical information (i.e., $(1 - H(P|V))\cdot H(P|V)$). In each coevolutionary scenario, the pollinator fitness is always optimized by the highest certainty of chemical information (i.e., low conditional entropy of pollinators to VOCs) provided by plant VOCs ($1 - H(V|Po)$) (Fig. 1). The optimal coevolutionary scenario was then determined by comparing the simulation results with the empirical data collected from fig–pollinator mutualisms, a paradigm of obligate pollination systems with a surprisingly long coevolutionary history having been initialized in the Late Cretaceous (c. 75 Ma) (Cook & Rasplus, 2003; Cruaud et al., 2012). Moreover, based on simulations, we evaluated the minimal number of VOCs that are essential for the plant–pollinator identification at the equilibrium of coevolution.

2 Material and Methods

2.1 Data collection and characterizing VOC profiles of receptive figs

In fig–pollinator mutualisms, a *Ficus* species is only pollinated by a single or only a few specific pollinating wasp species (Agaonidae), whose larvae can only develop inside the galled ovules of host figs (syconia). In addition to a large number of parasitic nonpollinating wasps (Agaonidae) that utilize figs by galling (Wang et al., 2019), many generalists from Lepidoptera, Diptera, and Coleoptera have also been observed to feed on figs (Miao et al., 2011). Host identification of pollinating wasps has been proved to rely on the VOCs emitted by the figs at the receptive stage (Chen et al., 2009; Souza et al., 2015). Although previous studies have obtained VOC profiles emitted by receptive figs from many *Ficus* species, whether these VOCs are either associated with pollinating wasps or utilized by antagonists to identify host plants is largely unclear. The exact components of attractants have been only detected in several fig–pollinator mutualisms, with very few (one to five) VOCs being functional to attract pollinating wasps (Chen et al., 2009; Zhang et al., 2020, Wang et al., 2021). This suggests that the majority of fig VOCs could be used to defend against antagonists (Volf et al., 2017) and serve as antimicrobials (Atkinson, 2018).

To screen the published data of VOC profiles of receptive figs, we searched in Web of Science (https://www.webofscience.com) using the combinations: (*Ficus* OR fig) AND (scents OR volatile OR chemical OR odors OR VOC). We only retained the publications where: (i) VOC collection was carried out for entire receptive figs using dynamic headspace technology, (ii) complete chromatograms of VOCs emitted from receptive figs were available using gas chromatography–mass spectrometry, and (iii) the composition of pollinating wasps was explicit for each studied *Ficus* species. Moreover, although intraspecific variation of VOCs can assist to explore the likely processes of microevolution leading to divergence of chemical communication modes among related species, the empirical VOC data were collected from only one or very few individuals in a single population for almost all the selected *Ficus* species (see Table S1). We therefore only focused on the interspecific variations of VOCs, and merged the VOC data from different individuals in the same population or haphazardly selected one population to form a single VOC profile for each selected *Ficus* species. We then checked the names of chemical compounds throughout the empirical data by searching the National Institute of Standard and Technology (NIST) database.

To characterize and compare profiles of VOCs among different *Ficus* species, we first described the composition of VOCs by allocating all detected VOCs into six groups of chemicals (aliphatics, benzenoids and phenylpropanoids, C5-branched compounds, terpenoids, miscellaneous cyclic compounds, and nitrogen-containing compounds) based on the classification proposed by Knudsen et al. (2006). We analyzed the accumulation curves of VOCs across all sampled species to assess whether our sampling efforts were sufficient. We then calculated the frequency of each VOC detected in the selected *Ficus* species (i.e., emergence frequency). The VOCs were further classified into one of six

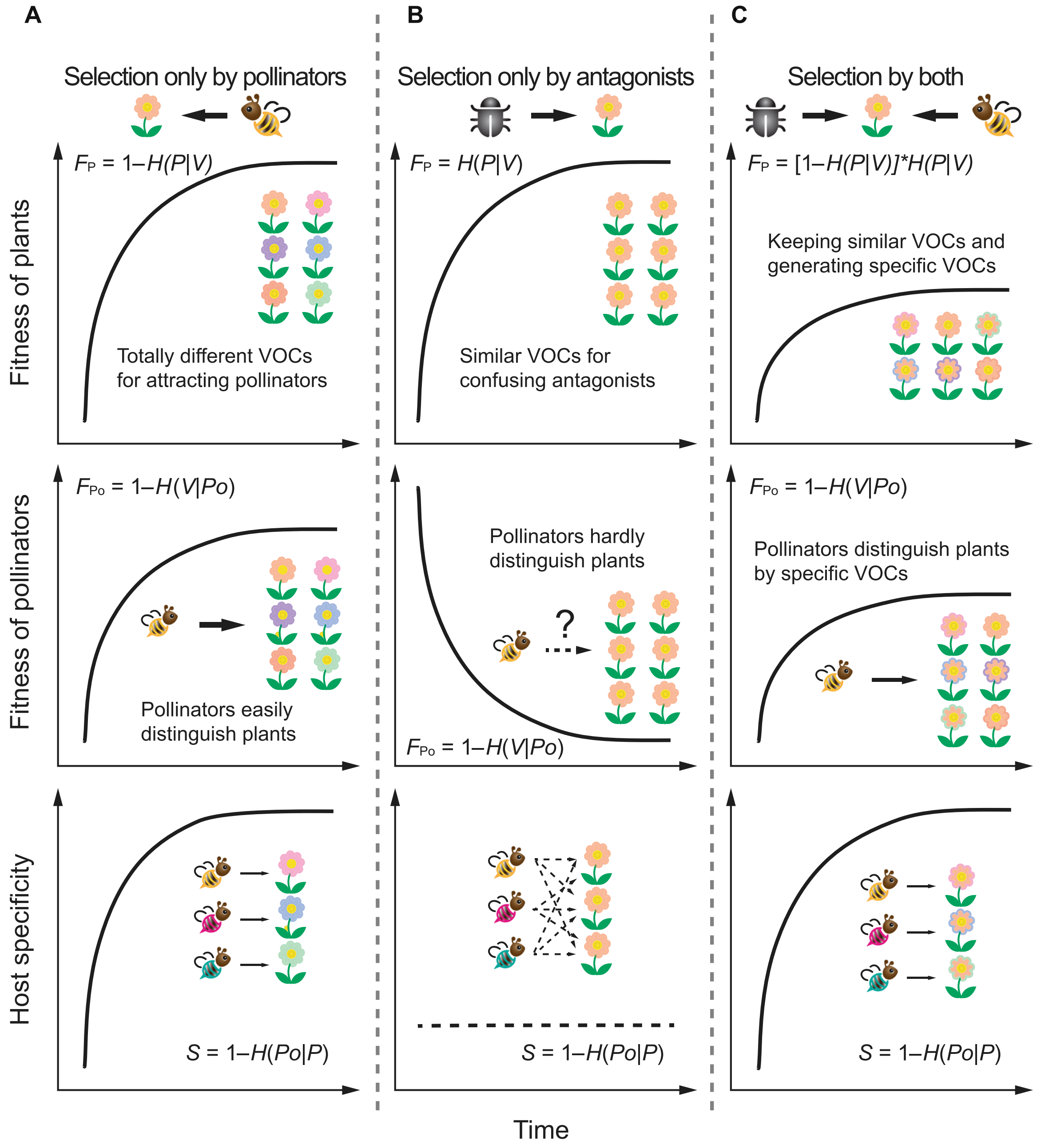

Fig. 1. Continued

categories of generalization, from emerging in less than 10% of species to very common: 0%–10%, 10%–20%, 20%–30%, 30%–40%, 40%–50%, and >50%. Meanwhile, we estimated the relative content of each VOC in each chromatogram by comparing the ion intensities of VOCs, which was assigned into one of the following six categories, from scarce to dominant: 0%–10%, 10%–20%, 20%–30%, 30%–40%, 40%–50%, and >50%. We could only carry out qualitative comparisons because absolute quantification was not available for the chromatograms obtained from different publications. To test the relationship between emergence frequencies and relative contents of VOCs, we used Spearman's correlation test setting with 10 000 permutations. We further examined whether the dominant or scarce VOCs were concentrated in some particular groups of chemicals using enrichment analysis in package *vegan* in R version 3.6.1 (R Core Team, 2019).

To evaluate the effects of phylogeny of *Ficus* species and their pollinating wasp species on the differentiation of VOC profiles, we calculated the Bray–Curtis distances among VOC profiles based on the approach in Salazar et al. (2018). We then tested the correlation between the Bray–Curtis distances and the phylogenetic distances among *Ficus* species/their pollinating wasp species using Mantel tests in package *vegan*, setting 10 000 permutations. Phylogenetic trees were constructed using the Bayesian algorithm by BEAST version 2.6.3 (Bouckaert et al., 2019), based on the sequences of five genes (internal transcribed sequence [*ITS*], external transcribed sequence [*ETS*], glyceraldehyde 3-phosphate dehydrogenase [*G3pdh*], chloroplast expressed glutamine synthetase region [*ncpGS*], and granule-bound starch synthase [*Waxy*]) for the plants, and six genes (elongation factor-la [*EF1a*], *Wingless* [*Wg*], cytochrome c oxidase subunit I [*COI*], cytochrome b [*Cyt b*], 28 S rRNA [D2–D3 and D4–D5 expansion regions], and 18 S rRNA [variable regions V3–V5]) for pollinating wasps. Phylogenetic distances were then calculated by MEGA version 5.1.0 (Tamura et al., 2011). All sequences were downloaded from GenBank (https://www.ncbi.nlm.nih.gov/genbank/).

2.2 Simulation of coevolution based on conditional entropy

To explore the optimal coevolution model explaining the properties of empirical VOC data, we conducted simulations under different coevolutionary scenarios between plants and pollinators, according to conditional entropy within the framework of information theory. Conditional entropy ($H_n(O|S) \in (0,1)$) shows the uncertainty of correctly identifying an object (O) when given a signal (S), which is calculated based on the object–signal (OS) matrix that has m objects and n signals (i.e., os_{ij} ($i \in [1,m]$, $j \in [1,n]$) represents an element in the matrix) using the formula proposed by Zu et al. (2020):

$$H_n(O|S) = -\sum_{j=1}^{n} P(S_i) H_n(O|S_j), \tag{1}$$

where $P(S_j) = \frac{\sum_{i=1}^{m} os_{i,j}}{\sum_{i=1,j=1}^{m,n} os_{i,j}}$ represents the probability of a given signal j being detected in all objects, and $H_n(O|S_j) = -\sum_{i=1}^{m} P(O_i|S_i) \log_n(O_i|S_j)$ reflects the uncertainty of correctly taking all figs apart by a given signal with $P(O_i|S_j) = \frac{os_{i,j}}{\sum_{i=1}^{m} os_{i,j}}$ corresponding to the probability of detecting object i given that signal j was observed. Given the concept of conditional entropy and that there were two groups of objects (plants and pollinators) in our study, we constructed two object–signal matrices PV (plants–VOCs) and PoV (pollinators–VOCs, which was derived from $PoP \times PV$, where the binary matrix PoP defined the plant–pollinator mutualisms). The size of each matrix was determined by the number of VOCs and fig–pollinator mutualisms from the empirical data, and no further data information was used in simulations (e.g., the composition of VOCs and detailed fig–pollinator associations). Because chromatogram data only provided relative contents of VOCs, which cannot be compared between species, we transformed the empirical data into binary data.

We undertook simulations to estimate the maximal fitness of both mutualists using conditional entropy under the following three coevolutionary scenarios (Fig. 1): (I) selection only by pollinators where the fitness of plants is expected to be $F_P = (1 - H(P|V))$, which means that the reduction in conditional entropy of PV will facilitate the attraction of specific pollinating wasps; (II) selection only by antagonists where $F_P = H(P|V)$, showing that plants emit VOCs mainly to confuse antagonists and the benefits from pollinators are negligible; and (III) selection by both pollinators and antagonists where $F_P = (1 - H(P|V)) \cdot H(P|V)$, meaning that the benefits from emitting explicitly distinguishable VOCs to attract pollinators $(1 - H(P|V))$ are at the cost of increasing risks of being attacked by antagonists ($H(P|V)$). In each of these three scenarios, pollinator fitness is considered to be $F_{Po} = 1 - H(V|Po)$, as lower conditional entropy of PoV indicates higher probability to identify a host plant species. Moreover, as $H(Po|P)$ reflects the uncertainty of plant–pollinator mutualisms, we define the host specificity of pollinators $S = 1 - H(Po|P)$, and increasing and decreasing value of S through evolution time means specialization and generalization, respectively (Fig. 1).

To evaluate the potential impacts of initial status on simulations, we set three different initial conditions in each coevolutionary scenario as proposed by Zu et al. (2020): (i) specialized interactions, in which each plant species only interacted with two pollinator species and six VOCs; (ii) random interactions, where a plant species randomly interacted with 50% and 75% of pollinator species/VOCs; and (iii) generalized interactions, where a plant species interacted

Fig. 1. Schematic representation of the hypothesized evolutionary trends of fitness of plants and pollinators and host specificity of pollinators under different plant–pollinator coevolutionary scenarios concerning chemical communications (I–III) based on conditional entropy. Plant fitness is determined by: (I) only the certainty of chemical information attracting pollinators (i.e., $1 - H(P|V)$, where $H(P|V)$ is the conditional entropy of plants to volatile organic compounds [VOCs]), (II) only the uncertainty of chemical information which confuses antagonists (i.e., $H(P|V)$), and (III) both aspects (i.e., $(1 - H(P|V)) \cdot H(P|V)$). Pollinator fitness is always optimized by the highest certainty of chemical information provided by plant VOCs (i.e., $1 - H(V|Po)$, where $H(V|Po)$ is the conditional entropy of pollinators to VOCs), and the host specificity of pollinators is defined as $S = 1 - H(Po|P)$, where $H(Po|P)$ is the conditional entropy of plant–pollinator relationships. Each column represents a coevolutionary scenario. In scenario I (column **A**) (selection on VOC composition only by pollinators), fitness of both plants and pollinators is expected to increase because a plant species tends to generate more specific VOCs to distinguish itself from other sympatric plants and attract pollinators regardless of the exposure to antagonists, and these specific VOCs are likely to establish more specialized mutualism with pollinators. In scenario II (column **B**) (VOC composition is selected only by antagonists), high similarity of VOCs is predicted to evolve to confuse antagonists irrespective of simultaneously confusing pollinators, leading to generalized mutualisms and increasing plant fitness but decreasing pollinator fitness. In scenario III (column **C**), where VOC composition is constrained by both pollinators and antagonists, plants share many VOCs for chemical defenses against antagonists but generate some specific VOCs to attract pollinators, and thus the fitness of both mutualists increases, and their mutualisms become specialized. Flowers with the same colors represent similar VOC profiles among plant species (**B**), and flowers with different colors indicate highly different VOC profiles among plant species (**A**) or VOC profiles containing many species-specific VOCs.

with all but one or maximum two pollinator species and VOCs. We then ran simulations using R version 3.6.1. For each initial condition, we calculated conditional entropy and fitness for both the plants and the pollinators. Then, in each loop, we mutated one element in each matrix (i.e., changing from 0 to 1 and from 1 to 0) each time and all abovementioned parameters were recalculated until 20% of total elements were mutated. If the new fitness of either mutualists was higher than the previous one, the mutated matrix would replace the previous matrix otherwise the previous matrix was kept for the next loop. In each simulation, we undertook 1000 loops to reach the equilibrium and obtain the maximal fitness.

To select the optimal coevolutionary scenario, we ran the simulation under each scenario 30 times and extracted conditional entropy and fitness of plants and pollinators at equilibrium (average values of the last 250 loops) from each run. Then we calculated the difference in conditional entropy and fitness between the simulated results and those from the empirical data (using the same formulas in the three coevolutionary scenarios). We chose the scenario owning the least difference as the best fitting scenario, with significance of this difference tested using ANOVA (Turkey's honestly significant difference [HSD] tests for pairwise comparisons) in R version 3.6.1.

To test the robustness of the optimal coevolutionary scenario, we first checked whether the simulation results were strongly impacted by different initial conditions and mutation rates (5% per loop). To examine whether our simulation results were stable with varying numbers of sampled species, we used a subset of empirical data only containing the fig–pollinator mutualisms in Eurasia to rescale the sizes of the three matrices and ran simulations with comparisons among different coevolutionary scenarios using the same approach as mentioned above.

2.3 Estimating the minimal number of VOCs for host identification using mutual information

To assess the minimal number of VOCs emitted by receptive figs that can provide sufficient information for host identification by pollinating wasps, we analyzed accumulative mutual information for both *Ficus* species and their pollinating wasps using the two *OS* matrices *PV* and *PoV*. Mutual information ($I(O,S) \in (0,1)$) reflects the amount of information for distinguishing objects by knowing the available signals (Shannon, 1948), which is calculated based on the formula proposed by Zu et al. (2020):

$$I(O, S) = H_m(O) - H_n(O|S), \tag{2}$$

where the information entropy of *m* objects is calculated using the formula $H_m(O) = -\sum_{i=1}^{m} P(O_i) log_n P(O_i)$, $P(Oi) = \frac{1}{m}$, and $H_n(O|S)$ is the conditional entropy.

To ensure that the study system was at equilibrium, we calculated the mutual information for both mutualists only using the final *PV* and *PoV* matrices in simulations under different coevolutionary scenarios. In each run of coevolutionary simulation, we calculated mutual information of plants for every combination of VOCs to determine the minimal number of VOCs that provides the maximal mutual information. For pollinators, as mutual information can only be analyzed by binary data, we transformed the final *PoV* matrix (which might be a nonbinary matrix because it was derived from the product of two binary matrices) in each coevolutionary scenario into 30 replicates of binary matrices to mimic the decoding process of pollinating wasps (see Zu et al., 2020). We then calculated the mutual information for pollinators in each replicate using the same method for plants and used the median value of the 30 replicates to represent the mutual information at a given number of VOCs. Finally, we compared the results from the empirical data with those from simulations to further confirm the optimal coevolutionary scenario driving the formation of coding–decoding strategy in fig–pollinator mutualisms.

3 Results

3.1 Data collection and characters of VOC profiles

After searching through Web of Science, we found 778 related publications, of which only 15 publications contained complete chromatograms of VOCs emitted from receptive figs. These chromatograms belonged to 45 *Ficus* species (5 subgenera and 11 sections) distributed in Eurasia, the Afrotropical region, and Australasia (Table S1). In addition, the composition of pollinating wasps had been well-studied for all these *Ficus* species (Table S1).

A total of 244 VOCs were identified from the chromatograms of receptive figs from the 45 *Ficus* species (Table S2). After allocating all the VOCs into different chemical groups, terpenoids (127) and aliphatics (60) were the most abundant chemical compounds (Fig. 2A). The results from accumulation curves showed that the number of VOCs gradually reached a plateau with increasing numbers of *Ficus* species, indicating sufficient sampling efforts in our data collection (Fig. S1A). We detected similar numbers of species-specific VOCs (116) and common VOCs (shared by at least two species) (128), and there was a significant positive correlation between the emergence frequencies and the relative contents of VOCs ($\rho = 0.288$, $p < 0.001$, Fig. S1B). Moreover, enrichment analysis revealed that the VOCs with higher relative contents were more likely to have higher emergence frequencies (Fig. 2B).

After searching GenBank, we obtained DNA sequences from 41 *Ficus* species and 38 pollinating wasp species for construction of phylogenetic trees and calculation of phylogenetic distances (Table S3; Fig. S2). There was no significant correlation between Bray–Curtis distances among VOC profiles and the phylogenetic distances among our selected *Ficus* species (Fig. 2C). In contrast, we found a significant association between Bray–Curtis distances among VOC profiles and the phylogenetic distances among pollinating wasps (Fig. 2D).

3.2 Determination of optimal coevolutionary scenarios

Based on the empirical data, we constructed a *PV* matrix (45×244), a *PoP* matrix (45×45), and a *PoV* matrix (45×244) for coevolutionary simulations. The results showed that: (i) in coevolutionary scenario I (selection only by pollinators), conditional entropy of the *PV*, *PoV*, and *PoP* matrix ($H(P|V)$, $H(Po|V)$, and $H(Po|P)$) at equilibrium of

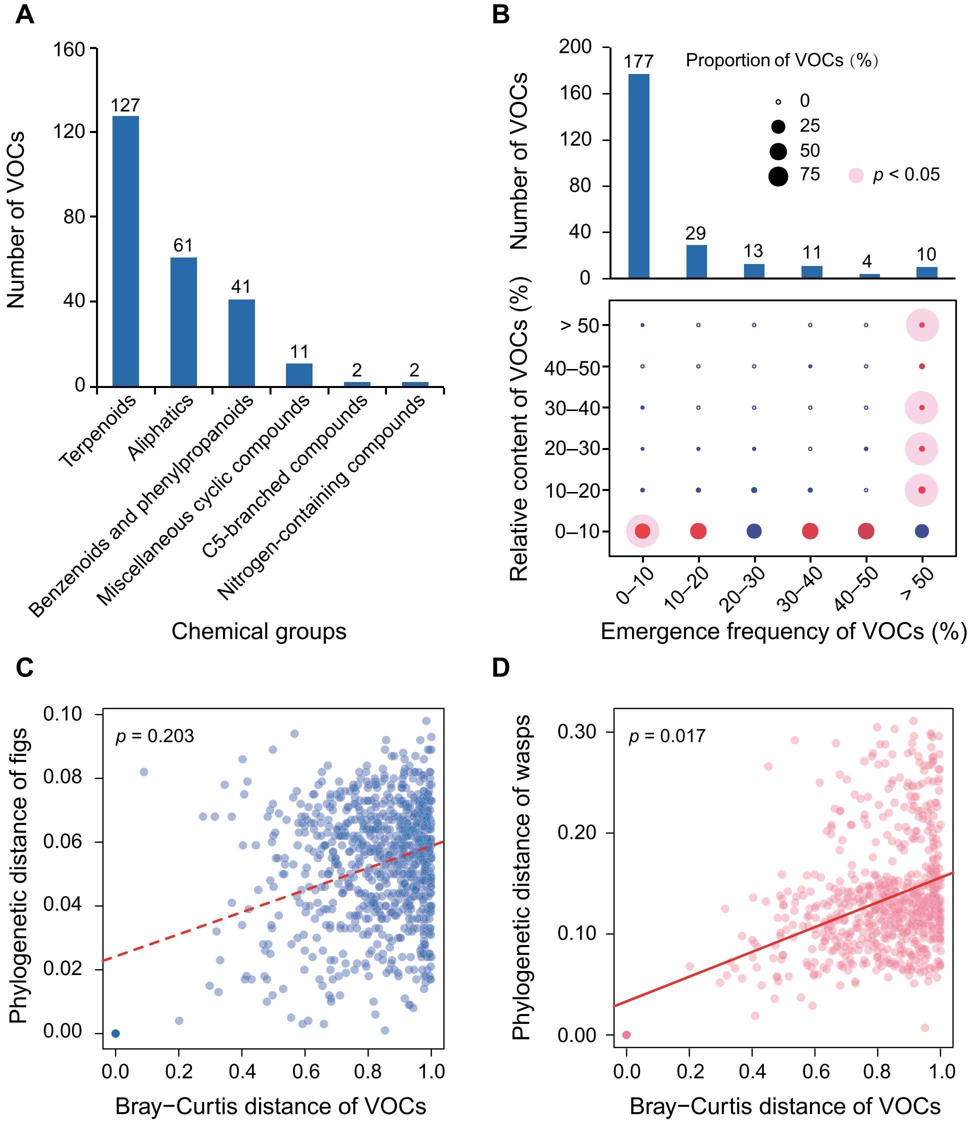

Fig. 2. Statistical properties of volatile organic compounds (VOCs) from receptive figs of *Ficus* species. **A**, VOC abundance in the six chemical groups. **B**, Numbers of VOCs in different categories of emergence frequency and relative content using enrichment analysis. **C**, Correlation between Bray–Curtis distances among VOC profiles and the phylogenetic distances among *Ficus* species. **D**, Correlation between Bray–Curtis distances among VOC profiles and the phylogenetic distances among pollinating wasps.

coevolution were 0.000, 0.066 ± 0.003, and 0.069 ± 0.003, respectively, with the maximal fitness of *Ficus* species (MF_P) and pollinating wasps (MF_{Po}) of 1.000 and 0.640 ± 0.002, respectively (Figs. S3A, S3B); (ii) in coevolutionary scenario II (selection only by antagonists), we found highly elevated values of $H(P|V)$ (1.000), $H(Po|V)$ (0.978 ± 0.001), and $H(Po|P)$ (0.514 ± 0.016) compared with those from the first scenario, leading to a high MF_P (1.000) but an extremely low MF_{Po} (0) (Figs. S3C, S3D); and (iii) in coevolutionary scenario III (selection by both pollinators and antagonists), relative to the above two scenarios, a medium level of $H(P|V)$ (0.500 ± 0.001) and $H(Po|V)$ (0.540 ± 0.001) and a low $H(Po|P)$ (0.061 ± 0.002) were detected, and MF_P and MF_{Po} reached 0.250 ± 0.001 and 0.320 ± 0.002, respectively (Figs. 3A, B). Furthermore, we found that different initial conditions did not change the final results of simulations under each coevolutionary scenario (Figs. 3A, 3B, S4A–S4D), exhibiting strong robustness in our simulations.

Using the empirical data, we calculated the conditional entropy and fitness of both mutualists, with $H(P|V) = 0.608$, $H(Po|V) = 0.607$, $H(Po|P) = 0.057$, $F_P = 0.238–0.608$, and $F_{Po} = 0.368$ (Table S4; Figs. 3C, 3D, S3E). Coevolutionary scenario III was the best fitting model for fig–pollinator mutualisms, because values of all the five parameters from this scenario showed the least differences with those from the empirical data (Figs. 3C, 3D), which were significantly smaller than the other two scenarios (Table 1). Under this scenario, simulation results showed a trend of specialization (increasing host specificity [S]) in plant–pollinator mutualisms, as $H(Po|P)$ decreased during the simulation process irrespective of initial

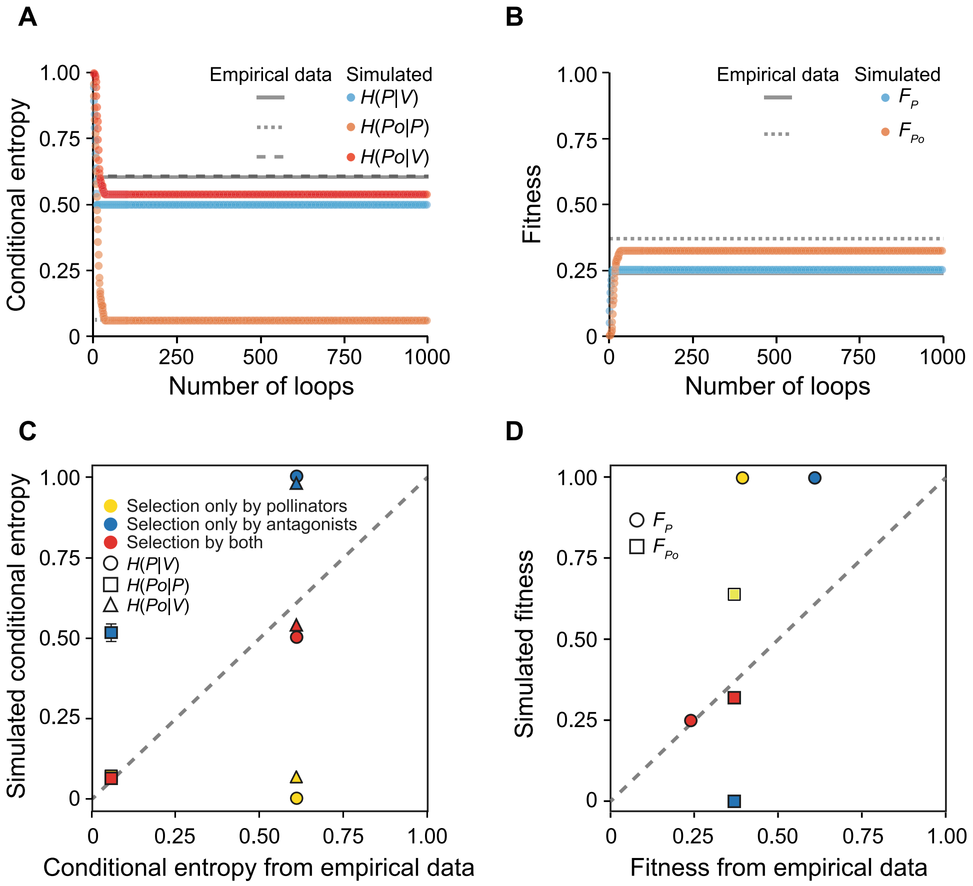

Fig. 3. Conditional entropy and fitness based on both coevolutionary simulations and empirical data. Conditional entropy (**A**) and fitness of both plants and pollinators (**B**) estimated by simulations under coevolutionary scenario III (selection by both pollinators and antagonists) and empirical data. Generalized interactions where a plant species interacted with all pollinator species and volatile organic compounds were set as the initial conditions of simulation with a mutation rate of 20% per loop. Similar results are obtained from different runs of simulation, and thus most data points at the same loop number are overlapped. Comparisons between (**C**) the conditional entropy at equilibrium and (**D**) the maximal fitness of both mutualists derived from simulations under the three coevolutionary scenarios and the conditional entropy/fitness based on the empirical data (Table S4; Fig. S3). The coevolutionary scenario with data points closest to the dashed line (*y*=*x*) shows the least difference with the empirical data and is considered as the optimal scenario.

conditions (from specialized to generalized interactions) (Figs. 3A, S4).

To further test the effects of varying number of sampled species, we used a subset of the empirical data only containing 25 *Ficus* species distributed in Eurasia (with 206 VOCs) and ran the simulations again. Coevolutionary scenario III was still the best fitting model and there was an evolutionary trend of specialization in plant–pollinator mutualisms (Tables S4, S5; Figs. S5A–S5E).

3.3 Estimation of the minimal number of VOCs for host identification

The accumulation curves of mutual information showed that: (i) in coevolutionary scenario I, mutual information of plants and pollinators reached to the maximum (i.e., 1) when one and two VOCs were randomly chosen (Fig. 4); (ii) in coevolutionary scenario II, we found very low mutual information (approaching zero) of both mutualists regardless of increasing number of VOCs (Fig. 4); and (iii) in coevolutionary scenario III, the minimal number of VOCs providing the maximal mutual information was estimated to be two for both mutualists (Fig. 4).

Based on the empirical data, we detected that the accumulation curves of mutual information reached to the maximum when only two VOCs were sampled for both mutualists, thus coevolutionary scenario III was again the best fitting model for fig–pollinator mutualisms (Fig. 4). When only using the data from Eurasian fig–pollinator mutualisms, the accumulation curves of mutual information were identical to those using the complete dataset (Figs. S5F, S5G).

4 Discussion

Chemical communications play a critical role in linking plants and their associated insects, and thus are expected to be shaped by their reciprocal adaptation (Whitehead & Peakall,

Table 1 Comparison of differences in conditional entropy and fitness between simulation results and empirical data among different coevolutionary scenarios using ANOVA with pairwise comparisons using Tukey's honestly significant difference (HSD) tests

	Difference between simulation results and empirical data (mean ± SE)					Pairwise comparison (Tukey's HSD)	
Parameter	Scenario I	Scenario II	Scenario III	*df*	*F*-value	Scenarios	Difference
$H(P \vert V)$	0.608 ± 0.000	0.392 ± 0.000	0.108 ± 0.000	2, 87	4.37×10^{10}***	Scenario I versus III	0.500***
						Scenario II versus III	0.284***
						Scenario I versus II	0.216***
$H(Po \vert V)$	0.542 ± 0.003	0.371 ± 0.001	0.069 ± 0.002	2, 87	10 811***	Scenario I versus III	0.473***
						Scenario II versus III	0.302***
						Scenario I versus II	0.171***
$H(Po \vert P)$	0.016 ± 0.002	0.457 ± 0.016	0.013 ± 0.002	2, 87	772***	Scenario I versus III	0.003^{NS}
						Scenario II versus III	0.444***
						Scenario I versus II	−0.441***
MF_P	0.608 ± 0.000	0.392 ± 0.000	0.012 ± 0.000	2, 87	4.43×10^{19}***	Scenario I versus III	0.596***
						Scenario II versus III	0.380***
						Scenario I versus II	0.216***
MF_{Po}	0.272 ± 0.002	0.368 ± 0.000	0.048 ± 0.002	2, 87	14 203***	Scenario I versus III	0.224***
						Scenario II versus III	0.320***
						Scenario I versus II	−0.096***

NS, not significant.
***$P < 0.001$.

2009). In this study, we performed simulations in the framework of information theory and found that the statistical properties of empirical VOC data of receptive figs were best explained by coevolutionary scenario III considering the selection from both attracting pollinators and confusing antagonists (Figs. 3C, 3D). Under such selection processes, fig–pollinator mutualisms became specialized, and as few as two VOCs could provide maximal mutual information for host identification of pollinators at the coevolutionary equilibrium. These findings support our hypothesis that conflicting selection could lead to specialized plant–pollinator mutualisms in chemical communications, which are in accordance with some previous studies (e.g., Zhang et al., 2020; Wang et al., 2021), where only a few attractive VOCs have been recorded for bridging a *Ficus* species and its pollinating wasps. In addition, our results showed a strong robustness (Fig. S5), and there were no significant footprints of *Ficus* phylogeny on empirical VOC composition (Fig. 2C).

The composition of VOCs reflects the trade-off between plant chemical defenses against antagonists and the requirement to attract pollinators, which seldom reaches to either extreme because plant species will lose their fitness if they only invest in chemical defenses (except for wind-pollinated plants) or pollinator attraction (Andrews et al., 2007; Schiestl et al., 2014). Meanwhile, the fitness of both pollinators and antagonists increases with more explicit chemical signals for host identification (Wiemer et al., 2009; Brodmann et al., 2012; Zu et al., 2020). Therefore, their decoding models (i.e., the range of VOCs from which they acquire information) could have diverged, otherwise plants can hardly benefit from attracting pollinators as the attractive VOCs also bring strong damages from antagonists. In contrast to most generalist antagonists, which have evolved to utilize many distinct VOCs to identify multiple host plants (Kessler et al., 2013; Segar et al., 2019), our results supported that pollinators might evolve to establish specific chemical communications with host plants like fig–pollinator mutualisms. Nevertheless, many nonpollinating wasps are genetically related to pollinating wasps (Wang et al., 2019), thus they might establish similar chemical communications with host figs to those between pollinating wasps and figs. Therefore, strong competition between pollinating and nonpollinating wasps are expected, which are also critical in the maintenance of fig–pollinator mutualisms and worth further exploration.

Such specialized mutualisms between pollinators and plants depending on only a few VOCs could take advantage of low uncertainty, in favor of their maintenance. This is further confirmed by evidence that maximal information for host identification of pollinators could be supplied by only two VOCs and that nearly half of VOCs in the empirical data are species-specific, consistent with some previous studies (Chen et al., 2009; Souto-Vilarós et al., 2018; Proffit et al., 2020). The relevant molecular mechanisms include the contraction of olfactory-related gene families in the pollinating wasps' genomes and the presence of a few genes regulating the biosynthesis of attractive VOCs in the plants (Zhang et al., 2020; Wang et al., 2021).

The evolutionary mechanisms forming such specialization could be relevant to niche differentiation in chemical communications between sympatric plant species by competition for pollinators and between allopatric species by adaptation to heterogenous environments (Segar et al., 2019). In either way, a certain level of intraspecific variations in VOC profile is required as the basis for the reciprocal adaptation between pollinators and plants, consequently altering chemical communication modes (Martinez-Swatson

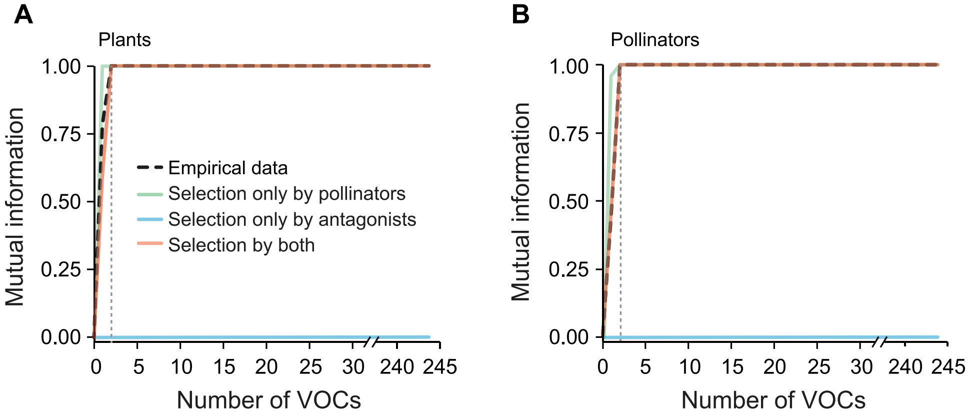

Fig. 4. Cumulative curves of mutual information for (**A**) plants and (**B**) pollinators with increasing number of given volatile organic compounds (VOCs) based on both simulation results under the three coevolutionary scenarios and the empirical data. In each coevolutionary scenario, data points at a given number of VOCs are overlapped because similar results are obtained from different runs of simulation.

et al., 2020). In addition to the genetic variations of plants, floral scents can also be influenced by microorganisms, probably to assist their spread by attracting pollinators (Cellini et al., 2019). However, to date, neither the geographic variations in VOCs nor the role of microorganisms in shaping the VOC profile has been comprehensively evaluated in *Ficus* species, thus we addressed the need for large-scale research on chemodiversity and microbes associated with *Ficus* species to improve our understanding of coevolution across multiple levels.

It is notable that our models only qualitatively utilized empirical data, without considering the concentration of each VOC. Although this disadvantage prevents the unraveling of the effects of mixtures of multiple VOCs on confusing antagonists, it might have limited impacts on pollinator host identification, because pollinating wasps can respond to the specific attractant even at extremely low concentration (Wang et al., 2021). In addition to the attractants, some other VOCs released by receptive figs might also participate in creating favorable atmospheres for pollinators. This is partly supported by the significant correlation between the phylogeny of pollinating wasps and the similarity of VOC profiles produced by their host *Ficus* species (Fig. 2D). Like the imprinting behavior of birds (Martinho & Kacelnik, 2016), many insects display associative learning (i.e., strong preferences to the VOC profiles in the places where they are born) (Proffit et al., 2015; Wang et al., 2017), hence imposing strong selection on VOC profiles.

Our results showed that plant–pollinator interactions can evolve towards specialization in chemical communications. This seems to contradict with the fact that there are numerous generalist pollinators (Ollerton, 2017). However, most of these generalists are defined by the number of their interacting species without the detailed knowledge about their evolutionary trajectories and the role of chemical communication in their interactions with plants (Mitchell et al., 2009; Zhang et al., 2020). Nowadays, biodiversity conservation is facing a major challenge of understanding the coevolutionary mechanisms of interacting species (Kiers et al., 2010; Potts et al., 2010; Segar et al., 2019). Future studies thus should pay more attention to unraveling the chemical communication modes of antagonists, and expand to specialized plant–pollinator–herbivore systems for further testing how chemical communications drive coevolution and diversification.

Acknowledgements

We thank Qian-Ya Li, Qi-Chong Zhu, Qiong Sun, and Shan Chen for their kind helps in collecting empirical data. This study was supported by the National Natural Science Foundation of China (31870356, 32171609, and 32071667).

Author Contributions

XYC, RW, and YY conceived and designed the study. YY collected empirical data and conducted data analyses and model simulations. RW, YYZ, and XYC edited the manuscript. All authors contributed to writing the manuscript.

Data Accessibility Statement

The R code for coevolutionary simulation will be available on request to the corresponding authors. All empirical data collected from 15 publications listed in Table S1.

References

Andrews ES, Theis N, Adler LS. 2007. Pollinator and herbivore attraction to *Cucurbita* floral volatiles. *Journal of Chemical Ecology* 33: 1682–1691.

Atkinson RG. 2018. Phenylpropenes: Occurrence, distribution, and biosynthesis in fruit. *Journal of Agricultural and Food Chemistry* 66: 2259–2272.

Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, De Maio N, Matschiner M, Mendes FK, Müller NF, Ogilvie HA, du Plessis L, Popinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu C-H, Xie D, Zhang C, Stadler T, Drummond AJ. 2019. BEAST

2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* 15(4): e1006650.

Brodmann J, Emer D, Ayasse M. 2012. Pollinator attraction of the wasp-flower Scrophularia umbrosa (Scrophulariaceae). *Plant Biology* 14: 500–505.

Cellini A, Giacomuzzi V, Donati I, Farneti B, Rodriguez-Estrada MT, Savioli S, Angeli S, Spinelli F. 2019. Pathogen-induced changes in floral scent may increase honeybee-mediated dispersal of *Erwinia amylovora*. *The ISME Journal* 13: 847–859.

Chen C, Song Q, Proffit M, Bessière J-M, Li Z, Hossaert-Mckey M. 2009. Private channel: A single unusual compound assures specific pollinator attraction in *Ficus semicordata*. *Functional Ecology* 23: 941–950.

Cook JM, Rasplus J-Y. 2003. Mutualists with attitude: Coevolving fig wasps and figs. *Trends in Ecology and Evolution* 18: 241–248.

Cruaud A, Rønsted N, Chantarasuwan B, Chou LS, Clement WL, Couloux A, Cousins B, Genson G, Harrison RD, Hanson PE, Hossaert-McKey M, Jabbour-Zahab R, Jousselin E, Kerdelhué C, Kjellberg F, Lopez-Vaamonde C, Peebles J, Peng YQ, Pereira RAS, Schramm T, Ubaidillah R, van Noort S, Weiblen GD, Yang DR, Yodpinyanee A, Libeskind-Hadas R, Cook JM, Rasplus J-Y, Savolainen V. 2012. An extreme case of plant–insect codiversification: Figs and fig pollinating wasps. *Systematic Biology* 61: 1029–1047.

Doyle LR. 2009. Quantification of information in a one-way plant-to-animal communication system. *Entropy* 11: 431–442.

Falara V, Amarasinghe R, Poldy J, Pichersky E, Barrow RA, Peakall R. 2013. The production of a key floral volatile is dependent on UV light in a sexually deceptive orchid. *Annals of Botany* 111: 21–30.

Johnson DS. 2010. The pollination niche and its role in the diversification and maintenance of the southern African flora. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365: 499–516.

Johnson MT, Campbell SA, Barrett SC. 2015. Evolutionary interactions between plant reproduction and defense against herbivores. *Annual Review of Ecology, Evolution, and Systematics* 46: 191–213.

Kessler A, Halitschke R. 2009. Testing the potential for conflicting selection on floral chemical traits by pollinators and herbivores: Predictions and case study. *Functional Ecology* 23: 901–912.

Kessler D, Diezel C, Clark DG, Colquhoun TA, Baldwin IT. 2013. *Petunia* flowers solve the defence/apparency dilemma of pollinator attraction by deploying complex floral blends. *Ecology Letters* 16: 299–306.

Kiers TE, Palmer TM, Ives AR, Bruno JF, Bronstein JL. 2010. Mutualisms in a changing world: An evolutionary perspective. *Ecology Letters* 13: 1459–1474.

Knudsen JT, Eriksson R, Gershenzon J. 2006. Diversity and distribution of floral scent. *The Botanical Review* 72: 1–120.

Lukas K, Harig T, Schulz S, Hadersdorfer J, Dötterl S. 2019. Flowers of European pear release common and uncommon volatiles that can be detected by honey bee pollinators. *Chemoecology* 29: 211–223.

Martinez-Swatson K, Kjøller R, Cozzi F, Simonsen HT, Rønsted N, Barnes C. 2020. Exploring evolutionary theories of plant defence investment using field populations of the deadly carrot. *Annals of Botany* 125: 737–750.

Martinho III A, Kacelnik A. 2016. Swapping mallards: Monocular imprints in ducklings are unavailable to the opposite eye. *Animal Behaviour* 122: 99–107.

McCall AC, Irwin RE. 2006. Florivory: The intersection of pollination and herbivory. *Ecology Letters* 9: 1351–1365.

Miao B-G, Yang D-R, Liu C, Peng Y-Q, Compton SG. 2011. The impact of a gall midge on the reproductive success of *Ficus benjamina*, a potentially invasive fig tree. *Biological Control* 59: 228–233.

Mitchell RJ, Irwin RE, Flanagan RJ, Karron JD. 2009. Ecology and evolution of plant–pollinator interactions. *Annals of Botany* 103: 1355–1363.

Ollerton J. 2017. Pollinator diversity: Distribution, ecological function, and conservation. *Annual Review of Ecology, Evolution, and Systematics* 48: 353–376.

Pichersky E, Raguso RA. 2018. Why do plants produce so many terpenoid compounds? *New Phytologist* 220: 692–702.

Potts SG, Biesmeijer JC, Kremen C, Neumann P, Schweiger O, Kunin WE. 2010. Global pollinator declines: Trends, impacts and drivers. *Trends in Ecology and Evolution* 25: 345–353.

Proffit M, Khallaf MA, Carrasco D, Larsson MC, Anderson P. 2015. 'Do you remember the first time?' Host plant preference in a moth is modulated by experiences during larval feeding and adult mating. *Ecology Letters* 18: 365–374.

Proffit M, Lapeyre B, Buatois B, Deng XX, Hossaert-Mckey M. 2020. Chemical signal is in the blend: Bases of plant–pollinator encounter in a highly specialized interaction. *Scientific Reports* 10: 10071.

R Core Team. 2019. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available from https://www.R-project.org/

Raguso RA. 2008. Wake up and smell the roses: The ecology and evolution of floral scent. *Annual Review of Ecology, Evolution, and Systematics* 39: 549–569.

Raguso RA. 2016. More lessons from linalool: Insights gained from a ubiquitous floral volatile. *Current Opinion in Plant Biology* 32: 31–36.

Raguso RA, Thompson JN, Campbell DR. 2015. Improving our chemistry: Challenges and opportunities in the interdisciplinary study of floral volatiles. *Natural Product Reports* 32: 893–903.

Salazar D, Lokvam J, Mesones I, Vásquez Pilco M, Ayarza Zuniga JM, de Valpine P, Fine PV. 2018. Origin and maintenance of chemical diversity in a species-rich tropical tree lineage. *Nature Ecology and Evolution* 2: 983–990.

Schiestl FP. 2015. Ecology and evolution of floral volatile-mediated information transfer in plants. *New Phytologist* 206: 571–577.

Schiestl FP, Kirk H, Bigler L, Cozzolino S, Desurmont GA. 2014. Herbivory and floral signaling: Phenotypic plasticity and trade-offs between reproduction and indirect defense. *New Phytologist* 203: 257–266.

Segar ST, Volf M, Sisol M, Pardikes NA, Souto-Vilarós D. 2019. Chemical cues and genetic divergence in insects on plants: Conceptual cross pollination between mutualistic and antagonistic systems. *Current Opinion in Insect Science* 32: 83–90.

Shannon CE. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27: 379–423.

Souza CD, Pereira RA, Marinho CR, Kjellberg F, Teixeira SP. 2015. Diversity of fig glands is associated with nursery mutualism in fig trees. *American Journal of Botany* 102: 1564–1577.

Souto-Vilarós D, Proffit M, Buatoism B, Rindos M, Sisol M, Kuyaiva T, Isua B, Michalek J, Darwell CT, Hossaert-McKey M, Weiblen GD, Novotny V, Segar ST. 2018. Pollination along an elevational gradient mediated both by floral scent and pollinator compatibility in the fig and fig–wasp mutualism. *Journal of Ecology* 106: 2256–2273.

Steffan-Dewenter I, Westphal C. 2008. The interplay of pollinator diversity, pollination services and landscape change. *Journal of Applied Ecology* 45: 737–741.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar, S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28: 2731–2739.

Vasiliev D, Greenwood S. 2020. Pollinator biodiversity and crop pollination in temperate ecosystems, implications for national pollinator conservation strategies: Mini review. *Science of the Total Environment* 774: 140880.

Volf M, Pyszko P, Abe T, Libra M, Kotásková N, Šigut M, Kumar R, Kaman O, Butterill PT, Sipos J, Abe H, Fukushima H, Drozd P, Kamata N, Murakami M, Novotny V. 2017. Phylogenetic composition of host plant communities drives plant–herbivore food web structure. *Journal of Animal Ecology* 86: 556–565.

Wang R, Chen XY, Chen Y, Wang G, Dunn DW, Quinnell RJ, Compton SG. 2019. Loss of top-down biotic interactions changes the relative benefits for obligate mutualists. *Proceedings of the Royal Society B: Biological Sciences* 286: 20182501.

Wang R, Yang Y, Jing Y, Segar ST, Zhang Y, Wang G, Chen J, Liu QF, Chen S, Chen Y, Cruaud Y, Ding YY, Dunn D, Gao Q, Gilmartin PM, Jiang K, Kjellberg F, Li HQ, Li YY, Liu JQ, Liu M, Machado CA, Ming R, Rasplus JY, Tong X, Wen P, Yang HM, Yang JJ, Yin Y, Zhang XT, Zhang YY, Yu H, Yue Z, Compton SG, Chen XY. 2021. Molecular mechanisms of mutualistic and antagonistic interactions in a plant–pollinator association. *Nature Ecology and Evolution* 5: 974–986.

Wang Y, Ma Y, Zhou DS, Gao SX, Zhao XC, Tang QB, Wang CZ, van Loon JJA. 2017. Higher plasticity in feeding preference of a generalist than a specialist: Experiments with two closely related *Helicoverpa* species. *Scientific Reports* 7: 17876.

Whitehead MR, Peakall R. 2009. Integrating floral scent, pollination ecology and population genetics. *Functional Ecology* 23: 863–874.

Wiemer AP, More M, Benitez-Vieyra S, Cocucci AA, Raguso RA, Sérsic AN. 2009. A simple floral fragrance and unusual osmophore structure in *Cyclopogon elatus* (Orchidaceae). *Plant Biology* 11: 506–514.

Zhang XT, Wang G, Zhang SC, Chen S, Wang YB, Wen P, Ma XK, Shi Y, Qi R, Yang Y, Liao ZY, Lin J, Lin JS, Xu XM, Chen XQ, Xu XD, Deng F, Zhao LH, Lee YL, Wang R, Chen XY, Lin YR, Zhang JS, Tang HB, Chen J, Ming R. 2020. Genomes of the banyan tree and pollinator wasp provide insights into fig–wasp coevolution. *Cell* 183: 875–889.

Zito P, Dötterl S, Sajeva M. 2015. Floral volatiles in a sapromyiophilous plant and their importance in attracting house fly pollinators. *Journal of Chemical Ecology* 41: 340–349.

Zu PJ, Boege K, Del-Val E, Schuman MC, Stevenson PC, Zaldivar-Riverón A, Saavedra S. 2020. Information arms race explains plant-herbivore chemical communication in ecological communities. *Science* 368: 1377–1381.

Zu PJ, García-García R, Schuman M, Saavedra S, Melian CJ. 2022. Plant–insect chemical communication in ecological communities: An information theory perspective. *Journal of Systematics and Evolution*. https://doi.org/10.1111/jse.12841.

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.12908/suppinfo:

Fig. S1. A, Accumulation curve of volatile organic compounds (VOCs) emitted by receptive figs of *Ficus* species. **B**, Correlation between the relative contents of VOCs and their emergence frequencies.

Fig. S2. Bayesian phylogenetic trees of (**A**) 41 *Ficus* species and (**B**) 38 pollinating wasp species.

Fig. S3. Simulation results of the conditional entropy at equilibrium and the maximal fitness of plants and pollinators under (**A, B**) coevolutionary scenario I (selection only by pollinator) and (**C, D**) coevolutionary scenario II (selection only by antagonists) and (**E**) the matrices of *PV*, *PoP*, and *PoV* with their sizes determined by the number of *Ficus* species, empirical volatile organic compounds (VOCs), and pollinating wasp species.

Fig. S4. Simulation results of conditional entropy and fitness using different initial conditions under coevolutionary scenario (**A, B**) I, (**C, D**) II, and (**E, F**) III.

Fig. S5. Simulation results based on the subset of empirical data only containing 25 *Ficus* species distributed in Eurasia.

Table S1. Basic information of the selected 45 *Ficus* species. The distribution ranges of these *Ficus* species were determined as described in Cruaud et al. (2012).

Table S2. Binary matrix of 244 volatile organic compounds (VOCs) detected in the chromatograms of receptive figs from 45 *Ficus* species.

Table S3. GenBank accession numbers (GAN) of DNA sequences for phylogeny reconstruction of 41 *Ficus* species and 38 pollinating wasp species.

Table S4. Conditional entropy and fitness of simulations under each coevolutionary scenario and the empirical data from both the 45 selected *Ficus* species and a subset of 25 *Ficus* species distributed in Eurasia.

Table S5. Comparison of differences in conditional entropy and fitness between simulation results and empirical data from the 25 *Ficus* species in Eurasia among different coevolutionary scenarios using ANOVA with pairwise comparisons using Tukey's honestly significant difference (HSD) tests.