Check for updates

Molecular mechanisms of mutualistic and antagonistic interactions in a plant-pollinator association

Rong Wang^{1,2,21}, Yang Yang^{1,21}, Yi Jing^{3,21}, Simon T. Segar^{4,21}, Yu Zhang¹, Gang Wang⁵, Jin Chen⁵, Qing-Feng Liu³, Shan Chen¹, Yan Chen⁶, Astrid Cruaud⁷, Yuan-Yuan Ding¹, Derek W. Dunn⁸, Qiang Gao³, Philip M. Gilmartin⁹, Kai Jiang¹, Finn Kjellberg¹⁰, Hong-Qing Li¹¹, Yuan-Yuan Li¹, Jian-Quan Liu¹², Min Liu¹³, Carlos A. Machado¹⁴, Ray Ming¹⁵, Jean-Yves Rasplus⁷, Xin Tong¹, Ping Wen⁵, Huan-Ming Yang³, Jing-Jun Yang¹, Ye Yin³, Xing-Tan Zhang¹⁶, Yuan-Ye Zhang¹⁷, Hui Yu^{18,19}, Zhen Yue³, Stephen G. Compton¹⁰²⁰ and Xiao-Yong Chen^{1,2}

Many insects metamorphose from antagonistic larvae into mutualistic adult pollinators, with reciprocal adaptation leading to specialized insect-plant associations. It remains unknown how such interactions are established at molecular level. Here we assemble high-quality genomes of a fig species, *Ficus pumila* var. *pumila*, and its specific pollinating wasp, *Wiebesia pumilae*. We combine multi-omics with validation experiments to reveal molecular mechanisms underlying this specialized interaction. In the plant, we identify the specific compound attracting pollinators and validate the function of several key genes regulating its biosynthesis. In the pollinator, we find a highly reduced number of odorant-binding protein genes and an odorant-binding protein mainly binding the attractant. During antagonistic interaction, we find similar chemical profiles and turnovers throughout the development of galled ovules and seeds, and a significant contraction of detoxification-related gene families in the pollinator. Our study identifies some key genes bridging coevolved mutualists, establishing expectations for more diffuse insect-pollinator systems.

volutionary adaptation fuels the genetic diversification of living organisms, driving speciation and emergent biodiversity^{1,2}. However, in contrast to adaptation to abiotic conditions^{3–5}, it remains unclear how species adapt to reciprocally evolving biotic factors at the molecular level. This reflects the difficulty of identifying the traits linking interspecific interactions in a dynamic selective landscape. The high diversity of phytophagous insects and angiosperms is believed to be the result of coevolution, in part driven by ongoing insect–plant arms races^{6,7}. Many herbivorous insects are also responsible for mediating gene flow between plant populations, often occurring as both antagonistic (that is, herbivorous) larvae and mutualistic pollinating adults⁸. Selection by multiple agents associated with herbivorous/pollinating insects acts on floral traits

both to deter herbivores and attract pollinators⁹, making it difficult to separate mechanistic processes in many plant–pollinator systems. Tightly coevolved species often have more apparent interacting traits, which provide an excellent testing ground for exploring coadaptation.

The obligate mutualisms comprising ~800 species from the genus *Ficus* (Moraceae) and their host-specific pollinating wasps (fig wasps; Hymenoptera, Agaonidae) form a classical example of coevolution and contribute greatly to ecosystem functioning, biodiversity and agriculture^{10,11}. Both mutualists have evolved strict correspondence in morphological, metabolic and life history traits^{10,12}. The plants reward the larvae of pollinating wasps with nutrition and protection, and each mutualist wasp species is both pollina-

¹Zhejiang Tiantong Forest Ecosystem National Observation and Research Station, Shanghai Key Lab for Urban Ecological Processes and Eco-Restoration, School of Ecological and Environmental Sciences, East China Normal University, Shanghai, China. ²Shanghai Institute of Pollution Control and Ecological Security, Shanghai, China. ³BGI Genomics, BGI-Shenzhen, Shenzhen, China. ⁴Agriculture and Environment Department, Harper Adams University, Newport, UK. ⁵CAS Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, China. ⁶Ecological Security and Protection Key Laboratory of Sichuan Province, Mianyang Normal University, Mianyang, China. ⁷INRAE, UMR 1062 CBGP, Montferrier-sur-Lez, France. 8College of Life Sciences, Northwest University, Xi'an, China. 9Department of Biological and Marine Science, University of Hull, Hull, UK. 10CEFE, CNRS, University of Montpellier, Paul Valéry University Montpellier, EPHE, IRD, Montpellier, France. 11School of Life Sciences, East China Normal University, Shanghai, China. ¹²Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu, China. ¹³School of Life Sciences, Guangzhou University, Guangzhou, China. ¹⁴Department of Biology, University of Maryland, College Park, MD, USA. 15 Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA. 16 Center for Genomics and Biotechnology, Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, Key Laboratory of Genetics, Breeding and Multiple Utilization of Corps, Ministry of Education, Fujian Agriculture and Forestry University, Fuzhou, China. ¹⁷Key Laboratory of the Ministry of Education for Coastal and Wetland Ecosystems, College of the Environment and Ecology, Xiamen University, Xiamen, China. ¹⁸Key Laboratory of Plant Resource Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China.¹⁹School of Life Sciences, Qufu Normal University, Qufu, China.²⁰School of Biology, University of Leeds, Leeds, UK. 21 These authors contributed equally: Rong Wang, Yang Yang, Yi Jing, Simon T. Segar. 🖾 e-mail: yuhui@scib.ac.cn; yuezhen@bgi.com; S.G.A.Compton@leeds.ac.uk; xychen@des.ecnu.edu.cn

ARTICLES



Fig. 1 | **Fig-pollinator mutualism between** *F. pumila* **var.** *pumila* **and Wiebesia** *pumilae*, **and determination of the compound attracting** *W. pumilae*. **a**, Life cycle of *W. pumilae* based on four fig developmental stages (pre-receptive, receptive, post-receptive and mature stages). This *Ficus* species is dioecious with figs on female trees growing long-styled female florets (seed florets) that are not available for pollinator oviposition. Therefore, female trees only produce seeds, while figs on functional male trees contain both male florets and short-styled female florets (feeder florets) that can be used by female pollinators for oviposition to support the larvae of the pollinators. At the receptive stage, adult female pollinators are attracted by host-specific VOCs and enter figs only through the ostiole (lined with bracts), either ovipositing into ovules of feeder florets in functional male figs or pollinating seed florets inside female figs. Pollinator larvae develop in induced galled ovules, and both larvae and seeds grow during the post-receptive stage. At the mature stage, after mating with adult males, adult female pollinators leave their natal figs carrying pollen donated by mature male florets and search for receptive figs and complete the cycle. **b**, Electrophysiological responses of a single female pollinator. N, nonanal; D, decanal; L, linalool. **c**, Electrophysiological responses of adult female pollinator. N, nonanal; D, decanal; L, linalool. **c**, Electrophysiological responses of adult female pollinator. N, nonanal; D, decanal; L, linalool. **c**, Electrophysiological responses of adult female pollinator. N, nonanal; D, decanal; L, linalool. **c**, Steve the sputhesized standard of each tentative COC compound (each electroantennogram curve represents five overlapped replicates). **d**, Preference of adult female pollinators to different tentative compounds using Y-tube olfactometer tests (NS, *P* > 0.05; ****P* < 0.001) (Supplementary Table 9).

tor and herbivore¹². Each individual wasp spends the majority of its lifespan at the larval stage (from three weeks up to nine months) and develops inside a single galled ovule of a female floret located inside the enclosed inflorescences characteristic of the genus (figs or 'syconia'13-15; Fig. 1a). There are two predominant types of breeding system in *Ficus* species, monoecy and dioecy¹⁶. In monoecious figs, each fig produces female florets that can be either pollinated or galled by pollinator larvae. In dioecious species, only the female florets (feeder florets) in figs of functional male trees support the development of pollinator offspring; figs growing on female trees attract pollinators to fertilize the female florets (seed florets) that do not support wasp development (Fig. 1a). On locating host figs, adult female wasps must crawl through a narrow passage usually lined by bracts (the ostiole) into a dark central lumen, where they typically remain trapped following oviposition and/or pollination. Short-lived (usually shorter than three days) adult wasps do not feed¹⁰.

Central to mediating these species-specific interactions are plant-emitted volatile organic compounds (VOCs), which guide adult female wasps to precisely identify and locate host figs^{13,14,17-21}. Moreover, the high-quality genomes of a *Ficus* species and its pollinating wasp (*Ficus microcarpa* and *Eupristina verticillata*) have been recently reported¹¹, which create a basis for exploring how these pollinators identify host figs at the molecular level. However, to date the key attractive VOCs have been explicitly identified in only a small number of fig–pollinator mutualisms^{17,18}, and the underlying molecular mechanisms determining host-specific signalling and insect attraction remain unknown.

Once the problem of host identification has been overcome, pollinator larvae must also survive and develop under a set of unique conditions inside galled ovules that support their development (Fig. 1a). While figs can defend against herbivores from a wide range of taxa²², it is unclear how pollinator larvae cope with plant defensive chemicals inside the galled ovules during this antagonistic phase of mutualism. One possible explanation is that galling behaviour may activate the plant reproductive program in galled tissues, so that galling insects can avoid the strong chemical defences induced by stress reaction when they utilize plant nutrients²³⁻²⁵. To test whether the reproductive program is activated in galled ovules, it is necessary to compare between the chemical profiles of galled ovules and seeds. We also expect that such adaptation to a specialized environment must leave molecular footprints in the pollinator genome, for example, contracted detoxification-related gene families^{26,27}.

Here we focused on a fig-pollinator mutualism comprising a dioecious *Ficus* species *Ficus pumila* var. *pumila*²⁸ and its specific pollinator *Wiebesia pumilae*²⁹ (Fig. 1a). We used multi-omics in combination with validation experiments to unravel the key molecular mechanisms contributing to the antagonistic and the mutualistic interactions in this system. We determined the specific attractive VOC and several key genes relevant to its biosynthesis. We identified the corresponding responses in the odorant-binding protein (OBP) genes in the pollinator genome and an OBP mainly binding

Table 1 | Summary statistics for the assembly of F. pumila var. pumila and W. pumilae genomes

Chromosome ID	F. pumila	var. pumila	W. pumilae	
	No. genes	Length (bp)	No. genes	Length (bp)
Chr1	1,697	20,463,500	816	21,315,831
Chr2	1,871	21,202,951	2,076	59,985,216
Chr3	2,335	23,199,346	2,631	66,440,284
Chr4	2,412	23,721,380	2,225	54,409,331
Chr5	3,327	31,603,922	2,281	59,419,729
Chr6	1,649	20,816,579	2,263	55,968,755
Chr7	2,070	23,331,000		
Chr8	2,097	21,006,959		
Chr9	1,856	21,788,500		
Chr10	1,740	20,107,953		
Chr11	1,798	22,360,920		
Chr12	2,000	20,847,995		
Chr13	2,526	34,592,857		
Number of contigs	543		102	
Total length of contigs (Mb)	315.7		318.2	
Contig N50 (Mb)	2.3		10.9	
Anchored genome content (Mb)	304.8		317.5	
Anchored rate (%)	96.6		99.8	
Scaffold N50 (Mb)	22.4		59.4	
Number of genes	28,187		12,316	

the attractant. During the antagonistic phase, we found a similar turnover of secondary metabolites when female florets developed into either galled ovules or seeds, and almost identical chemical profiles between these two tissues. It implies that the galled ovules may develop like seeds. A contraction of detoxification-related gene families was found in the pollinator genome, providing insights into the fig-pollinator coadaptation during antagonistic interaction.

Results

Assembly of genomes and evolution. To provide high-quality reference genomes for transcriptomic and proteomic analyses, we assembled genomes of F. pumila var. pumila and W. pumilae using a combination of Illumina and PacBio sequencing technologies (Supplementary Table 1 and Methods). The assembled genomes were 315.7 Mb (contig N50 of 2.3 Mb) for the plant and 318.2 Mb (contig N50 of 10.9 Mb) for the pollinator (Table 1 and Supplementary Table 2). Using the uniquely mapped reads produced by the high-throughput chromatin conformation capture (Hi-C) technique (Supplementary Tables 1 and 2), we generated Hi-C-based physical maps composed of 13 and 6 pseudo-chromosomes, with 96.6% (305 Mb) and 99.8% (318 Mb) of the assembled genomes anchored to the pseudo-chromosomes (Table 1 and Supplementary Fig. 1). The scaffold N50 of the assembled genomes were 22.4 Mb and 59.4 Mb, and the pseudo-chromosomes included 97.1% (27,378) and 99.8% (12,292) of protein-coding genes (Table 1). Genome annotation results showed that the structures and functions of 25,905 and 12,305 protein-coding genes were annotated in the two genomes (Supplementary Fig. 2 and Supplementary Tables 3-5). BUSCO quality analysis of annotation showed that 92.4% of

1,375 conserved plant genes and 91.3% of 4,415 Hymenoptera genes have complete coverage (Supplementary Table 3).

The protein-coding genes of F. pumila var. pumila and W. pumilae were clustered into 15,631 and 7,969 gene families (Supplementary Fig. 3 and Supplementary Table 6). Analysis of comparative genomics using the genomes of 13 angiosperm species including F. pumila var. pumila and three congeneric species (F. hispida11, F. microcarpa¹¹ and F. carica¹⁶) showed that in the common ancestors of the four Ficus species, 1,473 gene families had contracted and 888 gene families had expanded. Phylogenetic reconstruction revealed that F. hispida is more closely related to F. pumila var. pumila than the other two Ficus species (Supplementary Fig. 4a). In the analysis of comparative genomics using the genomes of 11 arthropod species containing W. pumilae and two other pollinator wasp species (Ceratosolen solmsi¹¹ and Eupristina verticillata²⁷), we found 48 expanded and 1,261 contracted gene families in the common ancestors of three pollinating wasp species. We recovered a group containing E. verticillata and W. pumilae with C. solmsi as its sibling (Supplementary Fig. 4b). There was no evidence for recent whole-genome duplication in the plant, and only a few small segments (total length of 1.3 Mb) were found to be duplicated in the pollinator genome (Supplementary Fig. 5).

Attractive compound forming fig-pollinator identification. At the receptive stage, figs release VOCs containing critical compound(s) attracting their pollinating wasps (Fig. 1a). To determine the attractive compound(s), we collected VOCs from functional male and female figs of F. pumila var. pumila at the pre-repetitive and the receptive stages using the dynamic headspace sampling (DHS) approach, and identified a total of 70 compounds (Fig. 1a, and Supplementary Tables 7 and 8). Only three (linalool, nonanal and decanal) of these compounds were found to elicit physiological responses of adult females of *W. pumilae* (Fig. 1b,c), of which only decanal was emitted exclusively at the receptive stage (Supplementary Table 8). We then conducted behavioural preference tests among the three compounds using 50 female pollinating wasps in each testing group. The wasps showed a significantly greater preference for decanal than the control and a significantly reduced preference for nonanal than the control, with a similar preference between decanal and a nonanal-decanal blend (Fig. 1d and Supplementary Table 9). These results demonstrate that the VOC compound decanal, emitted by F. pumila var. pumila figs at the receptive stage, functions to attract the pollinating wasp W. pumilae.

Molecular mechanisms of specific host identification. To identify the molecular mechanisms underlying the behavioural responses of W. pumilae to the VOCs emitted by its host figs, we annotated the four gene families involved in insect olfaction³⁰. Across these gene families, W. pumilae, E. verticillata and C. solmsi consistently have lower numbers of genes, and, in particular, the number of OBP genes is significantly lower than less host-specific insects (Fig. 2a). Phylogenetic and synteny analysis including genomes of the three pollinating wasp species and the distantly related Nasonia vitripennis showed that most OBP genes in the pollinating wasp species displayed strong homology and that the small number of OBP genes resulted from gene loss and infrequent tandem duplication (Supplementary Figs. 6a and 7a). There were apparent differences in motif structure among OBPs in six of the ten syntenic blocks (Supplementary Fig. 6b). The general contraction in OBP genes and frequent changes in motif structure of homologous OBPs among pollinating wasp species may be expected, given their high host specificity and different VOC cues used for detecting host figs.

Among the 12 OBP genes of *W. pumilae*, transcriptome and proteome evidence showed that all genes were transcribed but only seven are translated into detectable proteins in adult females (Fig. 2b, Supplementary Fig. 8 and Supplementary Table 7). There

ARTICLES



Fig. 2 | Molecular mechanisms of the specific host identification of *W. pumilae*. **a**, Number of genes in the four olfactory-related gene families (OBPs, ORs, CSPs and IRs) in different insect species. Significantly contracted families (***P < 0.001) are shown for *W. pumilae* and *C. solmsi*, and species are ranked according to their phylogeny (Supplementary Fig. 4b). **b**, Transcription and translation of OBP genes of adult females of *W. pumilae* not contacting (as the control) and contacting the VOCs emitted by *F. pumila* var. *pumila* figs at the receptive stage (Supplementary Table 10). FPKM, fragments per kilobase of exon model per million mapped fragments. **c**, Motif analysis predicting the most likely *W. pumilae* OBPs that can bind to decanal and nonanal (Supplementary Fig. 9). **d**, K_D values of the predicted OBPs to decanal and nonanal using SPR experiments (Supplementary Fig. 10 and Supplementary Table 11). Lower K_D indicates higher binding affinity, and error bars represent standard errors calculated by parameter estimation in the steady state affinity model ('M' represents mol/L).

were no proteins with significant differences in quantity (PSDs) and differentially expressed genes between the control and the VOCs-contacting treatment (Supplementary Table 10).

To explore functions of W. pumilae OBPs, we predicted motif structures of OBPs and compared them with the OBPs in Adelphocoris lineolatus³¹ and Culex quinquefasciatus³², known to have decanal or nonanal binding activity. Among the seven OBPs with detectable protein products, WpumOBP2 shows similar structure to the known decanal-binding protein and WpumOBP11 is similar to the known nonanal-binding protein (Fig. 2c and Supplementary Fig. 9). To validate the functions of WpumOBP2 and WpumOBP11, we produced the recombinant proteins for these two OBPs and measured their binding affinity to decanal and nonanal using surface plasmon resonance (SPR) experiments. Consistent with the prediction, the experiments revealed considerably lower binding affinity $(K_{\rm D};$ representing much higher binding affinity) of WpumOPB2 to decanal than to nonanal and far lower $K_{\rm D}$ of WpumOBP11 to nonanal than to decanal, and thus demonstrate the high binding affinity of these two OBPs to the corresponding compounds (Fig. 2d, Supplementary Fig. 10 and Supplementary Table 11). Therefore, these results provide solid evidence that WpumOBP2 is the main binding protein to the attractant, and pollination of F. pumila var.

pumila by *W. pumilae* is initiated by the binding of decanal with WpumOPB2.

Regulation of gene expression in attractant biosynthesis. To identify the tissue for attractant emission within figs, we measured the concentration of decanal emitted by ostiolar tissues and female florets at the receptive stage from both sexes of *F. pumila* var. *pumila* (Fig. 1a) using DHS, as previous studies from other species suggested VOCs are mainly released from these tissues^{13,20}. The concentration of collected decanal in ostiolar tissues was $3.13 \pm 1.11 \text{ pgg}^{-1}$ (mean ± s.e.), which was 9.1 times that in female florets ($0.34 \pm 0.05 \text{ pgg}^{-1}$; pairwise *t*-test: d.f. = 9, *t* = 6.02, *P* = 0.002). Thus, the results revealed that decanal was predominantly emitted by ostiolar tissues at a similar concentration between sexes (*t*-test: d.f. = 4, *t* = 0.20, *P* = 0.858 in ostiolar tissues; d.f. = 4, *t* = 0.24, *P* = 0.826 in female florets).

To identify key genes involved in the biosynthesis of decanal, we conducted transcriptome and proteome analyses on ostiolar tissues collected at the pre-receptive and the receptive stages (Supplementary Table 7). The biosynthesis of decanal and nonanal is involved in the pathways of fatty acid biosynthesis (ko00061), elongation (ko00062) and metabolism (ko00071 and ko00592; Fig. 3a)^{33,34}. Genes in these

pathways showed similar patterns of expression between transcriptome and proteome data (Supplementary Fig. 11 and Supplementary Table 12). Comparing the receptive with the pre-receptive stage, we detected a total of eight PSDs (Fig. 3b), probably facilitating the biosynthesis of decanal and suppressing the biosynthesis of nonanal at the receptive stage (Fig. 3a). Down-regulated PSDs included two long-chain acyl-CoA synthetases (ACSLs) and one very-long-chain (3R)-3-hydroxyacyl-CoA dehydratase (HACD), while up-regulated PSDs comprised an acetaldehyde dehydrogenase (ALDH), an alcohol dehydrogenase (ADH), two lipoxygenases (LOX2Ss) and one hydroperoxide lyase (HPL; Fig. 3b). To validate the function of key genes (the two ACSLs, the ALDH and the ADH) in decanal biosynthesis, we produced the recombinant proteins of these genes and conducted in vitro enzyme activity assay (see Methods). The final products of the in vitro reactions identified by liquid chromatography-mass spectrometry (LC-MS) or gas chromatography-mass spectrometry (GC–MS) are consistent with the standards (Fig. 3c-e). These results validate the enzyme activity of the two ASCLs in synthesizing hexadecanoyl-CoA as well as the ALDH and the ADH in synthesizing decanal and decanol.

To understand the transcriptional regulation of decanal biosynthesis, we conducted co-expression network analysis and found one module containing two key genes (FpumACSL10 and FpumALDH1) and four potential regulating transcription factors (two HD-ZIPs, one bHLH and one bZIP; Fig. 3f and Supplementary Table 13). Cis-element detection analysis revealed one G-box motif upstream of FpumACSL10 and six G-box and one HD-Zip motifs upstream of FpumALDH1 (Supplementary Table 14). As G-box binds to transcription factor families of bZIPs and bHLHs and HD-Zip binds to HD-ZIPs^{35,36}, we hypothesized that expression of *FpumACSL10* is regulated by the bHLH and the bZIP and all above four transcription factors regulate the expression of *FpumALDH1*. To test this hypothesis, we obtained gualified polyclonal antibodies for the four transcription factors and performed chromatin immunoprecipitation-quantitative polymerase chain reaction (ChIP-qPCR) experiments. High percentage input and fold enrichment values showed that the bHLH and the bZIP could bind to the promoter region of FpumACSL10 and all four transcription factors could bind to the promoter region of FpumALDH1 (Fig. 3g,h), providing evidence for our hypothesis.

Metabolic and genomic signature of antagonistic interaction. To understand the mechanisms of antagonistic interaction between figs of *F. pumila* var. *pumila* and larvae of *W. pumilae*, we analysed chemical profiles of different tissue types of female and functional male figs at the receptive and the post-receptive stage using metabolomic data (Supplementary Table 7 and Methods). We focused on the secondary metabolites associated with plant chemical defences (SMCDs)^{22,37–39}, comprising some terpenoids (triterpenes and sesquiterpenes) and phenylpropanoids (including their precursors and their derivatives; Supplementary Fig. 12). Metabolomic analysis revealed 736 SMCDs (108 terpenoids and 628 phenylpropanoids; Supplementary Table 15). While we found significant differences in chemical profiles between two types of tissue and between

different fig development stages, there were few differences between female and functional male figs (Fig. 4a). No secondary metabolites with significant difference in quantity (SMSDs) were found between feeder and seed florets at the receptive stage, and there were only three SMSDs between galled ovules and seeds at the post-receptive stage (Fig. 4b). Remarkably, we found similar changes of SMSDs in both the feeder floret–galled ovule and the seed floret–seed transitions (Fig. 4c and Supplementary Fig. 13a). Besides SMCDs, galled ovules and seeds shared similar overall chemical profiles (Supplementary Fig. 14). These results showed similar chemical changes and profiles in the development of female florets whether they were parasitized by pollinator larvae (becoming galled ovules) or not (developing into seeds).

As might be expected from organisms that spend most of their lives in a specific environment, contraction of three gene families crucial to the detoxification of plant defensive chemicals⁴⁰ (CYP450s, glutathione-S-transferases (GSTs) and carboxylesterases (CCEs) gene families) was found in the genomes of W. pumilae, E. verticillata and C. solmsi (Fig. 4d and Supplementary Fig. 15). Such contraction was mainly caused by gene loss and infrequent tandem duplication (Supplementary Fig. 7b-d), and most of the detoxification-related genes in the three pollinating wasp species were in the same monophyletic groups (Supplementary Fig. 16). Ten of the 56 detoxification-related genes in W. pumilae were at a high level (read counts >200) and were significantly up-regulated at the larval stage compared with the adult stage (Supplementary Fig. 17 and Supplementary Table 16). These metabolic and genomic signatures provide a molecular basis for further exploring the mechanisms of fig-pollinator coadaptation during their antagonistic interaction.

Discussion

Reciprocal selection on signalling and defence traits has shaped the molecular constraints governing how antagonistic larvae develop into mutualistic adult pollinators^{8,41}. In this study, our combination of classic electrophysiological experiments and multi-omics approaches has illuminated some key mechanisms forming the coadaptation in a pair of fig-pollinator mutualists. We identified the attractive VOC, detected that host identification by the specific pollinators may be linked to their reduced number of OBP genes, and validated an OBP mainly binding the attractant. We identified the key genes involved in the regulation of both attractant and repellent biosynthesis in the plant, from facilitating the synthesis of the repellent to favouring the production of the attractant. Surprisingly, matched changes in SMCDs occurred across the transitions from (1) floret to galled ovule and (2) floret to seed, and almost identical profiles of SMCDs were found in galled ovules and seeds. As for the pollinator, we detected a contraction of detoxification-related gene families.

Previous studies have mainly focused on the dominant components or the bouquet of components in the VOCs emitted by figs^{14,17}. In contrast, our results showed that only one VOC of relative low concentration attracts the focal pollinating wasp species,

Fig. 3 | **Regulation of gene expression in attractant biosynthesis in figs of** *F. pumila* **var. pumila. a**, Pathways associated with biosynthesis of decanal and nonanal (fatty acid biosynthesis (ko00061), elongation (ko00062) and metabolism (ko00071 and ko00592)). **b**, Fold changes of all PSDs and their transcriptomic expression between receptive and pre-receptive stages in ostiolar types in proteomes (Supplementary Table 12). NS, *P* > 0.05; **P* < 0.05; ***P* < 0.01; ****P* < 0.001. **c-e**, Results of in vitro functional characterization of the four key genes in the biosynthesis of decanal and nonanal. The peaks of synthesized standards and reaction products (treatments with enzyme added for three replicates) are shown for each key gene. Because there are two steps in the catalytic reaction of the two ACSLs, we show the ion intensity of both the intermediate product (hexadecanoyl-AMP) and the final product (hexadecanoyl-CoA) separated by LC-MS. The reaction products of the ALDH and the ADH (decanal and decanol) were identified using GC-MS. **f**, Transcriptomic expression of genes in the co-expression module including two key genes and the transcription factors predicted to regulate the expression of these two key genes (Supplementary Tables 13 and 14). 'F' and 'M' represent female and functional male figs. **g,h**, Results of ChIP-qPCRs (percentage input and fold enrichment) showing evidence that the predicted transcription factors can bind to the promoter regions of *FpumACSL10* (FPUM_023966-RA) and *FpumALDH1* (see Supplementary Tables 12 and 13). Error bars represent standard errors of experimental results.

ARTICLES

addressing the importance of detailing the complete spectrum of VOCs. Moreover, the attractive VOC (an aldehyde) in our focal species is distinct from the attractants found in other *Ficus* species

(usually terpenes)^{11,13,14,17,18}. Such a dramatic difference indicates deep divergence among congeners in the recognition of VOC attractants^{42,43}, providing the basis for adaptive radiations in both



NATURE ECOLOGY & EVOLUTION | VOL 5 | JULY 2021 | 974-986 | www.nature.com/natecolevol

NATURE ECOLOGY & EVOLUTION



Fig. 4 | Metabolic and genomic signature of antagonistic interaction between *F. pumila* **var**. *pumila* **and** *W. pumilae.* **a**, Results of PLS-DA for terpenoids (triterpenes and sesquiterpenes) and phenylpropanoids. Each oval indicates the 95% confidence intervals of a sample group. **b**, Distribution of SMCDs across fates of female florets. No SMSDs between feeder and seed florets and only three SMSDs (two down-regulated and one up-regulated) between galled ovules and seeds were found in the pathways related to plant chemical defences (Supplementary Fig. 12). **c**, Largely matched turnover of SMSDs in feeder floret-galled ovule and seed floret-seed transitions (*P* values are shown for Spearman's rank correlation tests). **d**, Number of genes in CYP450, CCE and GST gene families in different insect species. Significantly contracted families (****P* < 0.001) are shown for *W. pumilae* and *C. solmsi*.

Ficus and their pollinating wasps⁴⁴. In addition, similar concentration of the attractant emitted by different sexes of figs supports the intersexual mimicry hypothesis in *Ficus* species²¹, which argues that any changes in biosynthesis of attractant VOCs in female figs may cause loss of sexual reproduction²¹.

Similar chemical changes in the development of galled ovules and seeds and almost identical SMCD profiles in these tissues showed that the occupancy of pollinator larvae activates the reproductive program of galled ovules. This suggests that galling strategy may help pollinator larvae avoid the potential chemical sanctions when they exploit nutrients of host plants. This is likely to result from either pollinator larvae manipulating plant physiology or changes triggered by host figs once the feeder florets are galled. Chemical mimicry of fruits and seeds has been reported in other galling insects^{24,25}, while many studies also suggest that figs have evolved to accommodate pollinator larvae^{10,15}. Other possibilities, such as pollination before oviposition combined with minimal initial interference of pollinator larvae, can be largely excluded because most galled ovules were not pollinated. Furthermore, we collected figs at the middle (four weeks after the entrance of pollinator foundresses) of the post-receptive stage (generally lasting 8-10 weeks). Future research should perform bioassays to determine the chemicals inducing the development of galled ovules and the specific secondary metabolites defending against pollinator larvae. This will reveal how pollinator larvae activate the reproductive programs of host plants and why they can only utilize feeder florets.

The pollinating wasp species have evolved specializations in OBP and detoxification-related genes, probably because they are host specific and spend most of their lives inside galled ovules (although some detoxification-related genes are not only involved in detoxification but also important for the general life cycle of insets). Such specializations facilitate the maintenance of host specificity, but conservation of some OBP genes among pollinating wasps (Supplementary Figs. 6 and 7) may also offer opportunities for host shift^{19,45,46}. Moreover, selection to maximize pollinator fitness may drive rapid adaptive changes in fig traits such as floral scents, and such reciprocal selection has occurred in some generalized plant–pollinator systems⁴⁷.

Ongoing global changes are causing rapid evolution and phenotypic changes in many plants, leading to mismatches between key traits bridging plants and their pollinators^{4,48}. Erosion of these links can result in the collapse of long-evolved mutualisms and a loss of biodiversity, but may also lead to the rewiring of host association networks^{4,49,50}. Limitations to our knowledge of molecular determination in plant–pollinator interactions have made predictions about future changes in biodiversity and ecosystem functioning largely speculative. Our findings offer an example of gene for gene coadaptation that extends beyond the existing phenotype-based models of mutualism persistence⁵¹ and place trait-based multi-omics at the centre of the ecological and evolutionary research concerning interacting species in more diffuse systems.

Methods

Genome assembly and annotation. Genomic DNA was extracted from leaves of a female *F. pumila* var. *pumila* individual nearby Zhejiang Tiantong Forest Ecosystem National Observation and Research Station (TINAS; 121° 47′ E, 29° 48′ N), Ningbo, China, and from ~500 adult female pollinators of *W. pumilae* emerged from five figs on a functional male tree in South China Botanic Garden (SCBG; 113° 11′ E, 23° 11′ N), Guangzhou, China. Six pair-end and mate-pair libraries were prepared with varying insert sizes (Supplementary Table 1) for sequencing on an Illumina Hiseq 4000 platform. We also carried out PacBio Sequel platform. Based on the Illumina pair-end sequencing data, the genome sizes of both species were estimated by counting *k*-mer frequency using Jellyfish version 2.1.3⁵².

De novo genome assembly was conducted using MECAT version 1.2^{53} . The initial contig was polished twice based on raw PacBio data and then corrected twice using Illumina paired-end reads with Pilon version 1.22^{54} . Redundans version $0.13c^{55}$ was used to exclude redundant contigs, and we removed contaminative

sequences by searching against the NCBI nucleotide sequences database (ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/) using megablast⁵⁶ with *e* value $\leq 1 \times 10^{-5}$. Gap filling was implemented with PBJelly⁵⁷ after scaffolding based on Illumina mate-pair reads using BESST version 2.2.7⁵⁸.

To further improve the quality of genome assembly of both species, we used the Hi-C technique to scaffold contigs into pseudo-chromosomes. We constructed Hi-C libraries using the protocol described in ref. 59. Fresh leaves sampled from the same F. pumila var. pumila individual used in the above sequencing and adult female pollinators from SCBG were cross-linked by 4% formaldehyde solution, followed by an overnight digestion with a four-cutter restriction enzyme MboI (400 units) at 37°C, preparation DNA ends with biotin-14-dCTP and blunt-end ligation of the cross-linked fragments. Then, the proximal chromatin DNA was re-ligated by ligation enzyme, and the nuclear complexes were reverse cross-linked by proteinase K. After that, we extracted and purified DNA and removed biotin from non-ligated fragment ends using T4 DNA polymerase. The following steps included end reparation, enrichment of biotin-labelled Hi-C samples and ligation by Illumina paired-end sequencing adapters, and then the Hi-C library (insert size of 350 bp) was amplified by PCR and sequenced on an Illumina NovaSeq 6000 platform. High-quality data checked by HiC-Pro60 were mapped to genome using the Burrows-Wheeler-alignment tool (BWA)61, with extraction of uniquely mapped reads for pseudo-chromosome clustering and assembly using Juicer62 and 3D-DNA63.

Following genome assembly, we assessed completeness using BUSCO version 3.0.3⁶⁴ and Iso-Seq full-length transcripts. The high-quality full-length transcripts were mapped to genome assemblies using GMAP version 2014-12-21⁶⁵, setting a cutoff of aligned coverage at 0.85 and aligned identity at 0.9. The quality of genome assembly was further tested by mapping Illumina paired-end reads to the genome assemblies using BWA with the depth of coverage calculated using BamDeal version 0.19 (https://github.com/BGI-shenzhen/BamDeal/). For each species, Iso-Seq sequencing was performed using a PacBio Sequel platform, based on two SMRTbell libraries with insert sizes of 0-5 kb and 4.5–10 kb established by full-length complementary DNA. We used fresh leaves, young stems from fertile and sterile branchlets and figs at different developmental stages for the plant, and adult males and females for the pollinator.

Genome annotation includes repeat identification (including tandem repeats and transposable elements), annotation of non-coding RNAs, and gene prediction and annotation. When annotating repeat sequences, tandem repeats were identified using Tandem Repeats Finder version 4.07⁶⁶, and transposable elements were searched against Repbase 21.01⁶⁷ and the transposable element protein database using RepeatMasker version 4.0.6 (http://www.repeatmasker.org/) and RepeatProteinMask in RepeatMasker. LTR_Finder⁶⁸, PILER⁶⁹ and RepeatScout⁷⁰ were used to create a de novo transposable elements library, and the combined non-redundant library was classified by running RepeatMasker again.

To annotate non-coding RNAs, tRNAscan-SE version $1.3.1^{71}$ was used to identify transfer RNAs and their secondary structures. Small nuclear RNAs and microRNAs were searched for using INFERNAL version $1.1.1^{72}$ in the Rfam database version 12.0^{73} , followed by the detection of rRNAs by aligning with plant or invertebrate rRNA sequences using BLASTN (*e* value $\leq 1 \times 10^{-5}$).

Gene model prediction was conducted using the MAKER pipeline version 2.31.10⁻⁴. The Iso-Seq full-length transcripts, RNA-seq transcripts (assembled using Hisat2 version 2.0.1⁻⁵ and StringTie version 1.3.3⁻⁶), the protein sequences of related species and protein sequences from the Swiss-Prot database (https://www.uniprot.org) were included in the analysis. Ab initio gene prediction was performed with the gene predictors SNAP⁷⁷ and AUGUSTUS⁷⁸. The MAKER pipeline was run for two (for the plant) and three (for the pollinator) iterations for training and the final trained hidden Markov model was used for annotation. JBrowse version 1.12.3⁷⁹ was used to examine the gene models following each iteration. The gene models with the presence of a domain in Pfam database (http://pfam.xfam.org/) or with annotation edit distance (AED) ≤ 0.6 for *W. pumilae* and AED <1 for *F. pumila* var. *pumila* were retained. BUSCO was used to evaluate the completeness of gene annotation for both genomes.

After determining gene models, functions of protein-coding genes were annotated using DIAMOND version 0.8.23⁸⁰ by aligning them to the NCBI NR, Swiss-Prot⁸¹ and KEGG⁸² databases. Motifs and domains in protein sequences were annotated using InterProScan version 5.16–55.0⁸³ via searching public databases. Gene ontology terms for each gene were assigned using Blast2GO version 3.3⁸⁴.

Comparative genomics. To analyse the evolutionary characters of our studied genomes, we first carried out gene family clustering. The genome and annotation data of 13 other angiosperm species and 11 other arthropod species were downloaded (Supplementary Table 6). The gene models with open reading frames (ORFs) shorter than 90 bp were removed, and only the longest transcript was chosen to represent each gene. Gene family clustering was performed using OrthoMCL version 10–148⁸⁵ for the plant and TreeFam pipeline version 0.5.1⁸⁶ for the pollinator.

We then determined the phylogenetic relationships among the plants and among the insects in the species pools used in gene family clustering. Corresponding coding sequences were aligned based on the protein sequences of all single-copy orthologues using MUSCLE version 3.8.31⁸⁷, and codon position 2 of aligned corresponding coding sequences were concatenated into a super gene

NATURE ECOLOGY & EVOLUTION

using an in-house Perl script with a filtration of ambiguously aligned positions using trimAI version 1.4.1⁸⁸. After that, phylogenetic trees were reconstructed using PhyML version 3.0⁸⁹ using a general time reversible substitution model with a gamma distribution and 100 bootstrap replicates. PAML version 4.9⁹⁰ was used to estimate divergence time, setting 10,000 Markov chain Monte Carlo generations with a sampling frequency of 5,000 and a burn-in of 5 million iterations. Overall substitution rate was assessed using BASEML setting a REV substitution model.

Gene family expansion and contraction was analysed using CAFE version 2.1⁹¹, which employed a stochastic birth-and-death process to model the evolution of gene family sizes over a phylogeny. The birth-and-death parameter (λ) was estimated using 10,000 Monte Carlo random samples. We then used the family-wise method to statistically test if a gene family experienced significant expansion/contraction, and gene families with conditional *P* values less than 0.05 were considered to have accelerated rates of gains and losses.

We then tested whether the genomes of *F. pumila* var. *pumila* experienced whole-genome duplication. Syntenic blocks were identified using MCscan version 0.8⁹², and the rate of transversions on fourfold degenerate synonymous sites was calculated using the Hasegawa–Kishono–Yano substitution model to uncover potential speciation or whole-genome duplication events occurring in the evolutionary history of the plant. For *W. pumilae*, we tested for genomic segmental duplications. The self-alignment was performed using BLASTZ version 1.02⁹⁹, and a non-redundant set of segmental duplications was obtained using WGAC version 1.3⁹⁴.

Annotation of specific gene families and analysis of their evolution. To test whether the contraction specific gene families in W. pumilae, E. verticillata and C. solmsi contribute to the wasps' host-specificity and detoxification ability, we conducted a detailed annotation in chemosensory gene families (OBPs, chemosensory proteins (CSPs), olfactory receptors (ORs) and ionotropic receptor (IRs)) and detoxification gene families including CYP450s, GSTs and CCEs. The homologous genes of N. vitripennis, Apis mellifera and Drosophila melanogaster were used as queries to search the genome assembly of W. pumilae using TBLASTN at a criterion of *e* value $\leq 1 \times 10^{-5}$, and gene structures of identified genes were predicted using GeneWise version 2.4.195 with pseudogenes masked. We repeated this process iteratively until no more genes were detected. Additional genes from the MAKER annotation were also included if they included corresponding InterPro domains. All gene structures were manually checked and corrected if necessary, on the basis of full-length transcripts, RNA-seq transcripts and homologous proteins in JBrowse. We used the binomial distribution one-tailed test to examine gene family expansion/contraction among the compared species without considering their evolutionary relationships.

To reveal the evolutionary history of OBP, CYP450, CCE and GST gene families, syntenic blocks were identified across the genomes of the three pollinating wasp species and *N. vitripennis* using MCscan (https://github.com/tanghaibao/jcvi/wiki/MCscan-(Python-version)). We then constructed neighbour-joining phylogenetic trees to verify homologous genes among these insect species, using TreeBeST version 1.9.2% using a Jones–Taylor–Thornton model and 1,000 bootstraps.

VOC collection and component analysis. To reveal the composition of VOCs emitted by figs of *E punila* var. *punila* at different developmental stages, we collected VOCs from figs at both pre-receptive and receptive stages (Fig. 1a) in spring 2018 using DHS techniques⁵⁷. After a careful search, we chose ten mature *E punila* var. *punila* trees comprising five females and five functional males (Supplementary Table 8) nearby TINAS, within the natural range of the plant. Three figs were labelled on each selected individual. At each fig developmental stage (from early to middle April for pre-receptive stage and late April for receptive stage), we extracted the VOCs emitted by each labelled fig into an activated porapak adsorption tube (150 mg) during 8:00–11:00, using a protocol identical to ref.⁵⁷. Each adsorption tube was then eluted three times using 300µl *n*-hexane and stored at -20°C.

VOCs emitted by figs were then separated and identified using a coupled GC–MS system (HP 7890A-5975C, Agilent)⁹⁸. For each sample, 1.8 µl of eluate VOC extract, concentrated using nitrogen, was injected in split mode with a split ratio of 10:1 at 250 °C. Helium (1 ml min⁻¹) was used as carrier gas in a HP-5ms ($30 \text{ m} \times 250 \text{ µm} \times 0.25 \text{ µm}$, Agilent) GC column. We set the oven ramp at 40 °C for 1 min, and then 3 °C min⁻¹ to 140 °C for 1 min, followed by 5 °C min⁻¹ to 230 °C for 3 mins. Ionization was conducted by electron impact (70 eV, source temperature 230 °C). The MS quadrupole was heated to 150 °C, with the scanned mass range setting as from 40 to 550 mass-to-charge ratio (*m/z*). Compound identification was implemented by matching the mass spectra with NIST 08 MS libraries. We then calculated the relative proportions of all compounds emitted by figs at each developmental stage.

To evaluate the difference in the concentration of decanal between ostiolar tissues and female florets, we sampled figs at receptive stages from three female and three functional male individuals and identified the composition of VOCs emitted from these two types of tissue using the same approach mentioned above. The decanal concentration in each type of tissue in a plant individual was quantified by comparing its peak area with the internal standard (decyl acetate). **Electrophysiological responses of pollinating wasps.** To narrow the range of candidate VOCs attracting *W. pumilae*, we tested the electrophysiological responses of the pollinators to the collected VOCs, using gas chromatography–electroantennogram detection (GC–EAD). We used a system coupling a custom-made electroantennography⁶⁹ with a GC (Trace GC 2000, Themo Finnigan). We injected 1.8 µl of concentrated VOC extract eluate into the GC to separate different compounds. The GC conditions were identical to those used for the GC–MS component analysis, except that the oven ramp was set to 50 °C for 2 mins, and then to 10 °C min⁻¹ up to 280 °C for 1 min. After GC–flame ionization detector quantification, outflow from the GC column was delivered to the EAG as the stimulus through a custom, 40-cm-long heated (at 250 °C) transfer line with a clean, wet and static-free airflow. The stimulus was then puffed to the antenna of an adult female pollinator (collected from figs in TINAS) fixed onto the EAG with both ends of the antennae connected with prepared glass electrodes linking the probes of EAG to the potentiometric amplifiers.

This experiment was repeated five times (that is, antennae of five adult female pollinators), and the EAD signal was recorded using a HP 34465 A digital multimeter (Keysight). Both EAD and flame ionization detector signal data were aligned to verify the tentative compounds stimulating the adult female pollinator, using the software IO Libraries Suite 16 (Agilent) and BenchVue (Keysight). These tentative effective compounds were identified by matching the chromatographs with the results of component analysis using GC–MS.

We further tested the electrophysiological response of adult female pollinators (collected from figs in SCBG) to the synthesized standard of each tentative compound (Tokyo Chemical Industry, Japan; Toronto Research Chemicals, Canada), following the same procedures as above. A compound was determined as truly effective only when it was confirmed by the experiments using both eluate of VOC extracts and synthesized standard.

Behavioural preference of pollinating wasps. To test the behavioural preference of *W. pumilae* to different tentative effective VOCs, we used a Y-tube olfactometer (stem 8 cm, arms 9 cm, at an angle of 55°, internal diameter of 1.5 cm) following the methods described in ref.¹⁹. We placed the synthesized standard of each tentative effective VOC in the glass container, connecting one arm of the olfactometer to this treatment of *n*-hexane and blends of putative stimuli compounds and the other arm to the controls (only *n*-hexane; Supplementary Table 9). VOCs were passed from both arms to the stem through equal flow rates of cleaned and humidified airflow created by an air pump system with an activated charcoal filter and distilled water. To avoid visual distractions to the pollinators, we placed the olfactometer in the centre of a white table illuminated using three 40 W cool white fluorescent tubes above the arms.

Each healthy adult female pollinator (collected from figs in SCBG) was tested independently with an observation for 5 mins in the olfactometer, and its behaviour was assigned to one of the three choices: (1) towards the treatment (the insect went 1 cm past the Y junction (decision line) and stayed there more than 1 min); (2) towards the control; and (3) no choice (the insect did not reach the decision line within 5 mins). For each treatment–control combination, we repeated this experiment 50 times (that is, 50 adult female pollinators) and compared the proportions of different choices (towards the treatment and towards the control) using generalized linear models assuming binomial distribution of residuals to examine the preference of *W. pumilae*.

Sample collection for comparative transcriptome, proteome and metabolome. To reveal the molecular mechanisms forming the specific pollinator-host identification based on both transcriptomic and proteomic data, in spring 2017, after collecting several pre-receptive and receptive figs from the ten mature individuals of the plants used in VOCs collection (Supplementary Table 7), we dissected each sampled fig to gather ostiolar tissues with bracts. The total sample size therefore was 20 for the plant. In spring 2018, we sampled at least five figs at the mature stage from each of the five functional male mature individuals used for VOC collection (Supplementary Table 7). Each sampled mature fig was dissected into halves in a Teflon bag, and then each half was rapidly moved into a Teflon bag containing only clean air filtered by activated charcoal (as a control) or clean air and a receptive fig (as a treatment), to test whether differential expression occurred in some chemosensory genes when adult females were exposed to attractive VOCs. We then collected all adult females of W. pumilae emerging from the sampled figs according to the identity of functional male trees (a total of 10 samples with at least 100 adult female wasps in each sample). All sampled fig tissues and adult female pollinators were first stored in liquid nitrogen for 72 hours and then moved into a refrigerator at -80 °C.

To unravel how pollinator larvae adapt to the environments inside galled ovules using metabolomes, we sampled several receptive and post-receptive figs from the 10 plant individuals (Supplementary Table 7) and collected ostiolar tissues (20 samples), female florets (10 samples), galled ovules (5 samples) and seeds (5 samples) in spring 2020. For clearly distinguishing galled ovules and seeds from the female florets that were neither pollinated nor utilized by pollinators, the post-receptive figs were sampled four weeks after the entrance of adult female pollinators.

RNA-seq for *F. pumila* var. *pumila* and *W. pumilae*. After generating PCR-based libraries and sequencing on a BGISEQ500 platform (the Beijing Genomics

Institute, China), low-quality reads were filtered using SOAPnuke version 1.5.6¹⁰⁰. The acquired clean reads were then mapped to the genome assemblies of our studied species using Bowtie version $2.2.5^{101}$ and gene expression was quantified by RSEM version $1.2.12^{102}$.

Quantitative proteomes for F. pumila var. pumila and W. pumilae. We identified and quantified proteins for ostiolar tissues (sampled in 2017) using the isobaric tags for relative and absolute quantitation (iTRAQ)-based method. The strategy of quantifying proteomes was conducted according to the methods described in ref.¹⁰³. After total protein extraction, peptide labelling was performed using an iTRAQ Reagent 8PLEX kit according to the manufacturer's protocol. Extraction was followed by peptide fractionation, and the peptides separated from LC-20AD nano-HPLC (Shimadzu) were transferred into the tandem mass spectrometry Q-EXACTIVE (MS/MS; Thermo Fisher Scientific) for data-dependent acquisition (DDA) detection. After converting the raw MS/MS data into mascot generic format (MGF) using Proteome Discoverer version 1.2 (Thermo Fisher Scientific), the exported data in MGF format were searched using Mascot version 2.3 (Matrix Science) against the protein-coding sequences from our gene prediction. Quantification of proteins was achieved using IQuant¹⁰⁴, which uses the Mascot Percolator algorithm¹⁰⁵ to improve the results of peptide identification and the principle of parsimony to assemble proteomes. All the proteins with a false discovery rate (FDR)¹⁰⁶ of less than 1% were retained for further analyses.

We used a data-independent acquisition (DIA) approach to identify and quantify proteins in adult female pollinators (collected in 2018). Procedures identical to iTRAQ were first performed on the total protein extraction, peptide fractionation and peptides separation. Then, to create reference spectra for DIA, we first conducted DDA on a Q-EXACTIVE HF mass spectrometer (Thermo Fisher Scientific) coupled with an Ultimate 3000 RSLCnano system (Thermo Fisher Scientific) after a further peptide separation on an in-house packed nano-LC column (150 μ m × 30 cm, 1.8 μ m, 100 Å). Then, using the same instruments, DIA was performed following a brief procedure that consisted of a survey scan at 120,000 resolution from 400 to 1,250 m/z (maximal injection time (MIT) 50 ms), followed by scanning in DIA isolation windows setting 17 m/z with loop count 50 at 30,000 resolution (automatic gain control target 3×10^6 and auto MIT). The DDA spectra were identified by searching against the database of protein-coding sequences using the MaxQuant version 1.5.3.30107 at the FDR level of 1% with the minimum peptide length of 7. Based on the spectrogram database of DDA spectra, peptides and proteins in DIA data were identified and quantified using Spectronaut¹⁰⁸, employing the mProphet approach and setting iRT for retention time prediction. A target-decoy model was used to verify the quantification results at an FDR level of 1%.

Measurement of metabolomes of different types of tissue. Chromatographic separation of metabolites was performed using an ultra-performance liquid chromatography system (Waters), with an ACQUITY UPLC HSS T3 column $(100 \text{ mm} \times 2.1 \text{ mm}, 1.8 \mu\text{m}; \text{Waters})$ being used for the reversed phase separation and setting oven temperature at 50 °C and flow rate of 0.4 ml min⁻¹. After separation, gradient elution was conducted as follows: 0~2 min, 100% mobile phase A (water + 0.1% formic acid); 2~11 min, 0% to 100% mobile phase B (acetonitrile + 0.1% formic acid); 11~13 min, 100% B; 13~15 min, 0% to 100% A. The injection volume for each sample was 10 µl. Then, the eluted metabolites were identified in both positive and negative ion modes using a high-resolution tandem mass spectrometer Xevo G2 XS QTOF (Waters). The capillary and sampling cone voltages were set at 3.0 kV and 40.0 V for positive ion mode and at 2.0 kV and 40.0 V for negative ion mode. Mass spectrometry data were acquired in Centroid MSE mode, setting the time of flight mass range from 50 to 1,200 Da and the scan time of 0.2 s. For MS/MS detection, all precursors were fragmented at 20-40 eV with the scan time of 0.2 s. A quality control sample (pooling all samples together) was used after every 10 samples. Peak alignment, peak picking and quantitation of each metabolite were performed using Progenesis QI version 2.2, and the quality control based on LOESS signal correction¹⁰⁹ was conducted using quality control samples.

Comparative transcriptome, proteome and metabolome analysis. We carried out differential expression/concentration analysis for transcriptomes, proteomes and metabolomes, to detect the key genes contributing to the attractant-induced host specificity and the chemical cues for the adaptation of pollinator larvae to plant chemical defences.

For transcriptomes, differential expression was tested in ostiolar tissues between the pre-receptive and the receptive stage. For the pollinator, differential expression was conducted between contacting attractive VOC(s) versus not contacting, and between adults and larvae. We performed comparisons using DEseq2 version 1.4.5¹¹⁰ based on negative binomial distributions. *P* values were corrected using the Benjamini–Hochberg method for multiple tests. The differentially expressed genes with a fold change ≥ 2 and an adjusted *P* value ≤ 0.05 were considered statistically significant.

For proteomes, we tested the PSDs in ostiolar tissues between the pre-receptive and the receptive stages, and between sexes at each stage using IQuant, and PSDs were defined as fold changes in protein abundance ≥ 1.2 and Q value ≤ 0.05 . In

adult female pollinators, PSDs were analysed using MS stats^111 at criterions of fold changes ≥ 2 and Q value $\le 0.05.$

For metabolomes, to examine whether there were significant differences in the profile of SMCDs between different types of tissue and between the receptive and the post-receptive stages, we first carried out enrichment analysis to enrich all relevant secondary metabolites into the pathways associated with plant chemical defences, and then clustered all samples into different categories using the partial least squares-discriminant analysis (PLS–DA) model in metaX¹¹². Data were log₂-transformed and scaled by Pareto scaling. SMSDs were defined as variable importance for the projection calculated based on the first two axes from PLS–DA model ≥ 1 , fold change ≥ 1.2 or ≤ 0.83 and Q value ≤ 0.05 . In addition, we performed the PLS–DA model to test the difference in the entire profile of secondary metabolites between different types of tissue and between different fig developmental stages.

Motif analysis. We conducted motif analysis to check whether the OBPs in the same syntenic blocks among the three pollinating wasp species have similar motif structure using MEME Suite 5.0.4¹¹³. Motifs with *e* value \leq 0.05 were used for interspecific comparisons. To predict the most likely OBPs related to the identification of specific attractant and repellent, we created a dataset consisting of all OBPs in *W. pumilae* and the specific OBPs that had been verified to bind to decanal and nonanal in *Adelphocoris lineolatus*¹³ and *Culex quinquefasciatus*³⁴ as references, and performed motif analysis.

In vitro functional characterization of key genes. The full length of the ORF of the four key genes (for the plant) and of the two OBPs (for the pollinator; Supplementary Table 17) was confirmed by reverse transcription PCR and was then cloned into pET-28a (MilliporeSigma). After checking sequences by Sanger sequencing, these genes were expressed in *E. coli* strains BL21 (DE3) and Rosetta (DE3). The recombinant proteins produced were purified (purity >90%) using modified nickel-nitrilotriacetic acid agarose (Thermo Fisher Scientific).

We measured the affinities of the two OBPs to different substrates using SPR on a Biacore T200 system (GE Healthcare). OBPs were reconstituted in sterile phosphate buffered saline (PBS) and were diluted in 10 mM sodium acetate trihydrate (pH = 4.5) to the concentration of 20ug ml⁻¹. Then, each OBP was immobilized by the amine coupling method on a CM5 sensor chip according to the manufacturer's protocol (GE Healthcare). Analytes (decanal and nonanal) were diluted in running buffer (5% DMSO-PBS-P) to the concentration ranging from 0 to 1,000 μ M and were injected through channels at a flow rate of 20 μ l min⁻¹. Using BIAevaluation (GE Healthcare), both a steady state affinity model and a 1:1 binding model were performed to quantify $K_{\rm P}$.

For enzyme activity assays of the four key genes of the plant, we used the reaction system (500 μ l) mainly composed of 50 mM Tris-HCl (pH 7.4), 0.4–~1.0 mM substrate(s) (Supplementary Table 17), 2 M dimethyl sulfoxide (for the ADH and the ALDH)/10% triton X-100 (for the two ASCLs) and 10 μ l of purified protein (0.2 mg ml⁻¹). After 60 min of incubation at room temperature, we collected the reaction products by headspace solid-phase microextraction for the ALDH and the ADH (which were analysed by GC–MS) and by extraction using diethyl ether for the two ACSLs (which were analysed by LC–MS). These experiments were repeated three times. In addition, three replications of negative controls (only adding the subtrates and bovine serum albumin) were conducted, and no reaction products were detected.

Cis-element detection and co-expression network analysis. To test the regulatory mechanisms in the biosynthesis of attractant and repellent emitted by figs of *E. pumila* var. *pumila*, we first scanned the binding motifs present in the 2 kb promoter sequences upstream of key plant genes using PlantCARE¹¹⁴. Then, weighted undirected co-expression networks were conducted using the R package WGCNA¹¹⁵ with a soft thresholding power of 8. Modules containing genes with correlated expression patterns were identified by gene clustering based on the topological overlap matrix¹¹⁶ and by cutting the resulting dendrogram using the cutreeDynamic approach in the R package the Dynamic Tree Cut¹¹⁷. Genes with kME values larger than 0.95 were selected as hub genes. We checked whether some modules contained both some key plant genes and the transcription factors predicted to bind to them. This allowed us to uncover the likely regulatory mechanisms.

ChIP-qPCR. The ORF of each of the four transcription factors (*FpumHD-ZIP1*, *FpumHD-ZIP2*, *FpumbZIP1* and *FpumbHLH1*; Supplementary Table 13) was cloned into the pET-28a to generate the fusion plasmid encoding the six His-tagged fusion protein. This plasmid was transformed into *E. coli* strain Rosetta (DE3), which was cultured and induced by 0.8 mM isopropyl β -p-thiogalactoside at 37 °C. The induced cells were then sonicated for supernatant collection, and the purified recombinant proteins were obtained using a His-tag protein purification kit (Beyotime Biotechnology). The purified proteins were used to immunize rabbits for 52 days to acquire polyclonal antibody (ABclonal Biotechnology). We successfully obtained the qualified antibodies for all the four transcription factors for ChIP-qPCR experiments.

ChIP-qPCRs were then conducted for the two transcription factors with qualified antibodies to examine if they could bind the putative target genes

by model prediction. The ChIP assay was performed based on the protocols described in ref. ¹¹⁸. Approximately 3.0 g ostiolar tissues from figs at receptive stages were treated using 1% formaldehyde to crosslink and fix the DNA–protein complexes. The cells of sampled tissues were lysed, and each antibody was used to immunoprecipitate the antigen transcription factor with its binding DNA fragments. The DNA in the ChIP products was applied in qPCR with primer pairs designed for the promoters of putative target genes in a QuantStudio 5 real-time PCR detection system (Thermo Fisher Scientific). Each qPCR reaction was performed in triplicate and the cycle thresholds of ChIP products were compared with those of input samples and negative controls (only using immunoglobulin G (IgG)) for calculating percentage input and fold enrichment (% input (ChIP)/% input (negative control)). We failed to obtain the cycle threshold values for negative controls by the end of 35 qPCR cycles, and we therefore used the cycle threshold enrichment.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support the findings of this study have been deposited in the CNSA (https://db.cngb.org/cnsa/) of CNGBdb with accession code CNP0000674.

Code availability

All analyses in this study were conducted using published programs, and all codes for data analysis are provided in the Methods.

Received: 24 October 2019; Accepted: 20 April 2021; Published online: 17 May 2021

References

- 1. Lamichhaney, S. et al. Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* **518**, 371–375 (2015).
- 2. Simões, M. et al. The evolving theory of evolutionary radiations. *Trends Ecol. Evol.* **31**, 27–34 (2016).
- Arnegard, M. E. et al. Genetics of ecological divergence during speciation. *Nature* 511, 307–311 (2014).
- Hoffmann, A. A. & Sgrò, C. M. Climate change and evolutionary adaptation. *Nature* 470, 479–485 (2011).
- Olsen, J. L. et al. The genome of the seagrass *Zostera marina* reveals angiosperm adaptation to the sea. *Nature* 530, 331–335 (2016).
- Becerra, J. X., Nogeb, K. & Venable, D. L. Macroevolutionary chemical escalation in an ancient plant-herbivore arms race. *Proc. Natl Acad. Sci.* USA 106, 18062–18066 (2009).
- Edger, P. P. et al. The butterfly plant arms-race escalated by gene and genome duplications. *Proc. Natl Acad. Sci. USA* 112, 8362–8366 (2015).
- Adler, L. S. & Bronstein, J. L. Attracting antagonists: does floral nectar increase leaf herbivory? *Ecology* 85, 1519–1526 (2004).
- McCall, A. C. & Irwin, R. E. Florivory: the intersection of pollination and herbivory. *Ecol. Lett.* 9, 1351–1365 (2006).
- Cook, J. M. & Rasplus, J.-Y. Mutualists with attitude: coevolving fig wasps and figs. *Trends Ecol. Evol.* 18, 241–248 (2003).
- 11. Zhang, X. et al. Genomes of the banyan tree and pollinator wasp provide insights into fig-wasp coevolution. *Cell* **183**, 875–889 (2020).
- Herre, E. A., Jandér, K. C. & Machado, C. A. Evolutionary ecology of figs and their associates: recent progress and outstanding puzzles. *Annu. Rev. Ecol. Evol. Syst.* 39, 439–458 (2008).
- Souza, C. D. et al. Diversity of fig glands is associated with nursery mutualism in fig trees. Am. J. Bot. 102, 1564–1577 (2015).
- Souto-Vilarós, D. et al. Pollination along an elevational gradient mediated both by floral scent and pollinator compatibility in the fig and fig-wasp mutualism. *J. Ecol. Evol.* 106, 2256–2273 (2018).
- 15. Wang, R. et al. Loss of top-down biotic interactions changes the relative benefits for obligate mutualists. *Proc. R. Soc. B* **286**, 20182501 (2019).
- 16. Mori, K. et al. Identification of *RAN1* orthologue associated with sex determination through whole genome sequencing analysis in fig (*Ficus carica L.*). *Sci. Rep.* 7, 41124 (2017).
- 17. Proffit, M. et al. Chemical signal is in the blend: bases of plant-pollinator encounter in a highly specialized interaction. *Sci. Rep.* **10**, 10071 (2020).
- Chen, C. et al. Private channel: a single unusual compound assures specific pollinator attraction in *Ficus semicordata*. *Funct. Ecol.* 23, 941–950 (2009).
- Wang, G., Cannon, C. H. & Chen, J. Pollinator sharing and gene flow among closely related sympatric dioecious fig taxa. *Proc. R. Soc. B* 283, 20152963 (2016).
- Yu, H. et al. De novo transcriptome sequencing in *Ficus hirta* Vahl. (Moraceae) to investigate gene regulation involved in the biosynthesis of pollinator attracting volatiles. *Tree Genet. Genomes* 11, 91 (2015).

NATURE ECOLOGY & EVOLUTION

- Soler, C. C. L., Proffit, M., Bessière, J.-M., Hossaert -McKey, M. & Schatz, B. Evidence for intersexual chemical mimicry in a dioicous plant. *Ecol. Lett.* 15, 978–985 (2012).
- 22. Volf, M. et al. Community structure of insect herbivores is driven by conservatism, escalation and divergence of defensive traits in *Ficus. Ecol. Lett.* **21**, 83–92 (2018).
- Martinson, E. O., Hackett, J. D., Machado, C. A. & Arnold, A. E. Metatranscriptome analysis of fig flowers provides insights into potential mechanisms for mutualism stability and gall induction. *PLoS ONE* 10, e0130745 (2015).
- 24. Zhang, H. et al. Leaf-mining by *Phyllonorycter blancardella* reprograms the host–leaf transcriptome to modulate phytohormones associated with nutrient mobilization and plant defense. *J. Insect Physiol.* **84**, 114–127 (2016).
- Schultz, J. C., Edger, P. P., Body, M. & Appel, H. M. A galling insect activates plant reproductive programs during gall development. *Sci. Rep.* 9, 1833 (2019).
- The Nasonia Genome Working Group Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. *Science* 327, 343–348 (2010).
- 27. Xiao, J.-H. et al. Obligate mutualism within a host drives the extreme specialization of a fig wasp genome. *Genome Biol.* **14**, R141 (2013).
- Ohri, D. & Khoshoo, T. N. Nuclear DNA contents in the genus *Ficus* (Moraceae). *Plant Syst. Evol.* 156, 1–4 (1987).
- Chen, Y., Compton, S. G., Liu, M. & Chen, X.-Y. Fig trees at the northern limit of their range: the distributions of cryptic pollinators indicate multiple glacial refugia. *Mol. Ecol.* 21, 1687–1701 (2012).
- Chen, L.-G. et al. Binding affinity characterization of an antennae-enriched chemosensory protein from the white-backed planthopper, *Sogatella furcifera* (Horváth), with host plant volatiles. *Pestic. Biochem. Phys.* 152, 1–7 (2018).
- Gu, S.-H. et al. Functional characterization and immunolocalization of odorant binding protein 1 in the lucerne plant bug, *Adelphocoris lineolatus* (GOEZE). Arch. Insect Biochem. Physiol. 77, 81–99 (2011).
- Leal, W. S. et al. Reverse and conventional chemical ecology approaches for the development of oviposition attractants for *Culex* mosquitoes. *PLoS ONE* 3, e3045 (2008).
- Rizzo, W. B. et al. Fatty aldehyde and fatty alcohol metabolism: review and importance for epidermal structure and function. *Biochim. Biophys. Acta* 1841, 377–389 (2014).
- Schwab, W., Davidovich-Rikanati, R. & Lewinsohn, E. Biosynthesis of plant-derived flavor compounds. *Plant J.* 54, 712–732 (2008).
- Capella, M., Ribone, P. A., Arce, A. L. & Chan, R. L. Arabidopsis thaliana HomeoBox 1 (AtHB1), a Homedomain-Leucine Zipper I (HD-Zip I) transcription factor, is regulated by PHYTOCHROME-INTERACTING FACTOR 1 to promote hypocotyl elongation. *New Phytol.* 207, 669–682 (2015).
- Jiang, W. et al. Two transcription factors TaPpm1 and TaPpb1 co-regulate anthocyanin biosynthesis in purple pericarps of wheat. J. Exp. Bot. 69, 2555–2567 (2018).
- 37. Guan, R. et al. Draft genome of the living fossil *Ginkgo biloba*. *GigaScience* 5, 49 (2016).
- Mithöfer, A. & Boland, W. Plant defense against herbivores: chemical aspects. Annu. Rev. Plant Biol. 63, 431–450 (2012).
- Salazar, D. et al. Origin and maintenance of chemical diversity in a species-rich tropical tree lineage. *Nat. Ecol. Evol.* 2, 983–990 (2018).
- Després, L., David, J. P. & Gallet, C. The evolutionary ecology of insect resistance to plant chemicals. *Trends Ecol. Evol.* 22, 298–307 (2007).
- Segar, S. T., Volf, M., Sisol, M., Pardikes, N. & Souto-Vilarós, A. D. Chemical cues and genetic divergence in insects on plants: conceptual cross pollination between mutualistic and antagonistic systems. *Curr. Opin. Insect Sci.* 32, 83–90 (2019).
- 42. Cook, J. M. & Segar, S. T. Speciation in fig wasps. *Ecol. Entomol.* **35**, 54–66 (2010).
- 43. Cruaud, A. et al. An extreme case of plant–insect codiversification: figs and fig-pollinating wasps. *Syst. Biol.* **61**, 1029–1047 (2012).
- Yu, H. et al. Multiple parapatric pollinators have radiated across a continental fig tree displaying clinal genetic variation. *Mol. Ecol.* 28, 2391–2405 (2019).
- Satler, J. D. et al. Inferring processes of coevolutionary diversification in a community of Panamanian strangler figs and associated pollinating wasps. *Evolution* 73, 2295–2311 (2019).
- Wang, G. et al. Genomic evidence of prevalent hybridization throughout the evolutionary history of the fig-wasp pollination mutualism. *Nat. Commun.* 12, 718 (2021).
- 47. Hoballah, M. E. et al. Single gene-mediated shift in pollinator attraction in Petunia. *Plant Cell* **19**, 779–790 (2007).
- Potts, S. G. et al. Global pollinator declines: trends, impacts and drivers. Trends Ecol. Evol. 25, 345–353 (2010).

- Kiers, E. T., Palmer, T. M., Ives, A. R., Bruno, J. F. & Bronstein, J. L. Mutualisms in a changing world: an evolutionary perspective. *Ecol. Lett.* 13, 1459–1474 (2010).
- 50. Segar, S. T. et al. The role of evolution in shaping ecological networks. *Trends Ecol. Evol.* **35**, 454–466 (2020).
- Stoy, K. S., Gibson, A. K., Gerardo, N. M. & Morran, L. T. A need to consider the evolutionary genetics of host-symbiont mutualisms. *J. Evol. Biol.* 33, 1656–1668 (2020).
- Marçais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* 27, 764–770 (2011).
- Xiao, C.-L. et al. MECAT: fast mapping, error correction, and de novo assembly for single-molecule sequencing reads. *Nat. Methods* 14, 1072–1074 (2017).
- Walker, B. J. et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* 9, e112963 (2014).
- Pryszcz, L. P. & Gabaldón, T. Redundans: an assembly pipeline for highly heterozygous genomes. *Nucleic Acids Res.* 44, e113 (2016).
- Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. A greedy algorithm for aligning DNA sequences. J. Comput. Biol. 7, 203–214 (2000).
- English, A. C. et al. Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS ONE* 7, e47768 (2012).
- Sahlin, K., Chikhi & Arvestad, R. L. Assembly scaffolding with PE-contaminated mate-pair libraries. *Bioinformatics* 32, 1925–1932 (2016).
- 59. Belton, J. M. et al. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* 58, 268–276 (2012).
- Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* 16, 259 (2015).
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler Transform. *Bioinformatics* 25, 1754–1760 (2009).
- 62. Durand, N. C. et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).
- Dudchenko, O. et al. De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. Science 356, 92–95 (2017).
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212 (2015).
- Wu, T. D. & Watanabe, C. K. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* 21, 1859–1875 (2005).
- Benson, G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27, 573–580 (1999).
- 67. Bao, W., Kojima, K. K. & Kohany, O. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* **6**, 11 (2015).
- Xu, Z. & Wang, H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* 35, W265–W268 (2007).
- Edgar, R. C. & Myers, E. W. PILER: identification and classification of genomic repeats. *Bioinformatics* 21, i152–i158 (2005).
- Price, A. L., Jones, N. C. & Pevzner, P. A. De novo identification of repeat families in large genomes. *Bioinformatics* 21, i351–i358 (2005).
- Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25, 955–964 (1997).
- Nawrocki, E. P. & Eddy, S. R. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29, 2933–2935 (2013).
- 73. Nawrocki, E. P. et al. Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res.* **43**, D130–D137 (2015).
- Holt, C. & Yandell, M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinf.* 12, 491 (2011).
- Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360 (2015).
- Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295 (2015).
- Johnson, A. D. et al. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 24, 2938–2939 (2008).
- Stanke, M. & Morgenstern, B. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* 33, W465–W467 (2005).
- 79. Buels, R. et al. JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biol.* **17**, 66 (2016).
- Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60 (2015).
- Bairoch, A. & Apweiler, R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* 28, 45–48 (2000).

- Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44, D457–D462 (2016).
- 83. Jones, P. et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
- Conesa, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676 (2005).
- Li, L., Stoeckert, C. J. Jr & Roos, D. S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13, 2178–2189 (2003).
- 86. Li, H. et al. TreeFam: a curated database of phylogenetic trees of animal gene families. *Nucleic Acids Res.* **34**, D572–D580 (2006).
- 87. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
- Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973 (2009).
- Guindon, S., Delsuc, F., Dufayard, J. F. & Gascuel, O. Estimating maximum likelihood phylogenies with PhyML. *Methods Mol. Biol.* 537, 113–137 (2009).
- Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591 (2007).
- De Bie, T., Cristianini, N., Demuth, J. P. & Hahn, M. W. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* 22, 1269–1271 (2006).
- Tang, H. et al. Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps. *Genome Res.* 18, 1944–1954 (2008).
- Schwartz, S. et al. Human-mouse alignments with BLASTZ. Genome Res. 13, 103-107 (2003).
- Bailey, J. A., Yavor, A. M., Massa, H. F., Trask, B. J. & Eichler, E. E. Segmental duplications: organization and impact within the current human genome project assembly. *Genome Res.* 11, 1005–1017 (2001).
- Birney, E., Clamp, M. & Durbin, R. GeneWise and Genomewise. *Genome Res.* 14, 988–995 (2004).
- Ruan, J., Li, H., Chen, Z. & Coghlan, A. TreeFam: 2008 update. Nucleic Acids Res. 36, D735–D740 (2008).
- 97. Tholl, D. et al. Practical approaches to plant volatile analysis. *Plant J.* **45**, 540–560 (2006).
- Wen, B., Mei, Z., Zeng, C. & Liu, S. metaX: a flexible and comprehensive software for processing metabolomics data. *BMC Bioinform.* 18, 183 (2017).
- 99. Wen, P. et al. The sex pheromone of a globally invasive honey bee predator, the Asian eusocial hornet, *Vespa velutina*. *Sci. Rep.* **7**, 12956 (2017).
- Chen, Y. et al. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *Gigascience* 7, 1–6 (2018).
- Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359 (2012).
- Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinform.* 12, 323 (2011).
- 103. Tian, X., Chen, L., Wang, J., Qiao, J. & Zhang, W. Quantitative proteomics reveals dynamic responses of *Synechocystis* sp. PCC 6803 to next-generation biofuel butanol. *J. Proteom.* 78, 326–345 (2013).
- Wen, B. et al. IQuant: an automated pipeline for quantitative proteomics based upon isobaric tags. *Proteomics* 14, 2280–2285 (2014).
- Brosch, M., Yu, L., Hubbard, T. & Choudhary, J. Accurate and sensitive peptide identification with Mascot Percolator. *J. Proteome Res.* 8, 3176–3181 (2009).
- 106. Savitski, M. M., Wilhelm, M., Hahne, H., Kuster, B. & Bantscheff, M. A scalable approach for protein false discovery rate estimation in large proteomic data sets. *Mol. Cell Proteomics* 14, 2394–2404 (2015).
- 107. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372 (2008).
- Bruderer, R. et al. Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Mol. Cell Proteomics* 14, 1400–1410 (2015).
- 109. Dunn, W. B. et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* **6**, 1060–1083 (2011).
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550 (2014).
- Choi, M. et al. MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* 30, 2524–2526 (2014).
- 112. Wen, X.-L., Wen, P., Dahlsjö, C. A. L., Sillam-Dussès, D. & Šobotník, J. Breaking the cipher: ant eavesdropping on the variational trail pheromone of its termite prey. *Proc. R. Soc. B* 284, 20170121 (2017).

NATURE ECOLOGY & EVOLUTION

- Bailey, T. L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2, 28–36 (1994).
- 114. Lescot, M. et al. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **30**, 325–327 (2002).
- 115. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 559 (2008).
- 116. Yip, A. M. & Horvath, S. Gene network interconnectedness and the generalized topological overlap measure. *BMC Bioinform.* **8**, 22 (2007).
- Langfelder, P., Zhang, B. & Horvath, S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* 24, 719–720 (2008).
- Gendrel, A. V., Lippman, Z., Martienssen, R. & Colot, V. Profiling histone modification patterns in plants using genomic tiling microarrays. *Nat. Methods* 2, 213–218 (2005).

Acknowledgements

We thank Y.-J. Wang, Q.-C. Zhu, Q. Sun, Q.-Y. Li, J.-W. Wang, T.-L. Xu, Y. Chen, F.-L. Wei, G.-C. Shen, X.-L. Wen and L.-S. Li for their kind help in field experiments; X.-L. Wen, L.-S. Li and the Public Technology Service Center of XTBG (CAS) for assisting active VOCs analysis; and X.-G. Mao, P.-Y. Hua and D.-Y. Zhang for constructive suggestions in data analysis. This work is supported by NSFC grants 31630008 and 31870356 (X.-Y.C.) and 31870359 (G.W.), and a Talents 1000 Fellowship of Shaanxi Province (D.W.D.). S.T.S. acknowledges departmental support from Harper Adams University.

Author contributions

X.-Y.C. and R.W. conceived and designed the study. R.W., Y. Yang, S.G.C., S.T.S., H.Y. and Z.Y. conducted the experiments and analysed data. Y.J., Q.-F.L., H.Y., Y.-Y.Z., G.W., J.C., R.M., S.C., Y.C., D.W.D., H.-Q.L., M.L., Y.-Y.D., Y.-Y.L., X.T., P.W., J.-J.Y., X.-T.Z., Q.G., Y. Yin, K.J. and H.-M.Y. contributed to data acquisition and data analyses. R.W., S.T.S., S.G.C., J.-Q.L., J.-Y.R., F.K., C.A.M, A.C., P.M.G., Y.-Y.Z. and X.-Y.C. edited the manuscript. All authors contributed to writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-021-01469-1.

Correspondence and requests for materials should be addressed to H.Y., Z.Y., S.G.C. or X.-Y.C.

Peer review information *Nature Ecology & Evolution* thanks Shuqing Xu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021

nature research

Corresponding author(s): Xiao-Yong Chen

Hui Yu, Zhen Yue, Stephen G. Compton and Xiao-Yong Chen

Last updated by author(s): Apr 16, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our <u>Editorial Policies</u> and the <u>Editorial Policy Checklist</u>.

Statistics

Fora	all st	atistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.
n/a	Сог	nfirmed
	\boxtimes	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
	\boxtimes	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
	\boxtimes	The statistical test(s) used AND whether they are one- or two-sided Only common tests should be described solely by name; describe more complex techniques in the Methods section.
	\boxtimes	A description of all covariates tested
	\boxtimes	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
	\boxtimes	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
	\boxtimes	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted Give P values as exact values whenever suitable.
\boxtimes		For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
\boxtimes		For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
\boxtimes		Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
		Our web collection on <u>statistics for biologists</u> contains articles on many of the points above.

Software and code

Policy information about <u>availability of computer code</u>			
Data collection	No software was used, and all data were obtained from our experiments.		
Data analysis	All data analyses were conducted using programmes that have already been published, including Jellyfish version 2.1.3, MECAT version 1.2, Redundans version 0.13c, PBJelly, BESST version 2.2.7, BUSCO version 3.0.3, GMAP version 2014-12-21, BamDeal version 0.19, Tandem Repeats Finder (TRF) version 4.07, RepeatMasker version 4.0.6, LTR_Finder, PILER, RepeatScout, tRNAscan-SE version 1.3.1, INFERNAL version 1.1.1, BLAST, MAKER pipeline version 2.31.10, Hisat2 version 2.0.1, Stringtie version 1.3.3, SNAP, AUGUSTUS, Browse version 1.12.3, DIAMOND version 0.8.23, InterProScan version 5.16-55.0, Blast2GO version 3.3, OrthoMCL version 10-148, MUSCLE version 3.8.31, trimAl version 1.4.1, PhyML version 3.0, PAML MCMC version 4.9, CAFE version 2.1, MCscan version 0.8, BLASTZ version 1.02, WGAC version 1.3, TreeBeST version 1.9.2, SOAPnuke version 1.5.6, Bowtie version 2.2.5, RSEM version 1.2.12, Proteome Discoverer version 1.2, Mascot version 2.3, MaxQuant version 1.5.3.30, Spectronaut, DEseq2 version 1.4.5, IQuant, MSstats, R package WGCNA, and MEME Suite 5.0.4.		

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data that support the findings of this study have been deposited in the CNSA (https://db.cngb.org/cnsa/) of CNGBdb with accession code CNP0000674.

Field-specific reporting

Life sciences

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Behavioural & social sciences 🛛 Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see <u>nature.com/documents/nr-reporting-summary-flat.pdf</u>

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Many insects metamorphose from antagonistic larvae into mutualistic adult pollinators, with reciprocal adaptation at either stage leading to specialization of insect-plant associations. In this study, we combined omics data (de novo assemblies of genomes, transcriptomes, proteomes and metabolomes) from a fig-pollinator mutualism comprising Ficus pumila var. pumila and its host-specific pollinating fig wasp Wiebesia pumilae, to uncover some key molecular mechanisms contributing to the antagonism-mutualism transition of fig wasps and their impacts on host specificity. We identified a single volatile compound, decanal, as the 'private channel' attracting the host-specific pollinator wasp species of a single Ficus species. Based on omics data, we detected and validated a range of differentially expressed genes shaping the temporal- and tissue-specific emission of the attractant. Host identification of pollinators was found to rely on a highly-reduced number of odorant-binding protein (OBP) genes, and we also found an OBP mainly binding the attractant. In addition, similar chemical changes and profiles throughout the development of galled ovules and seeds showed that the plant reproductive program was activated in galled ovules. This matched with the contraction in several detoxification-related gene families in the pollinator genome. Our study provides an example molecular coadaptation in a tightly associated plant-pollinator mutualism, establishing expectations in more diffuse pollination systems.
Research sample	The mutualism between Ficus pumila var, pumila and its host-specific pollinating fig wasp Wiebesia pumilae typifies fig-pollinator
	mutualisms. We therefore selected this interacting pair as target species, to reveal the critical molecular mechanisms relevant to their reciprocal adaptation. In our study, we chose five functionally male (whose figs support the larvae of W. pumilae) and five female plants of F. pumila var. pumila (seed figs) as the experimental plants, to collect the two types of fig tissues as well as the VOCs emitted from the figs at both prereceptive and receptive stages (see methods and Supplementary Table 7), for the analysis of comparative transcriptome and proteome, and tests of behavior preference of pollinators. In addition, we sampled some individuals of W. pumilae from each of the five experimental functionally male plants (see methods and Supplementary Table 7) for comparative transcriptome and proteome analysis. Moreover, high-quality genomes of the two target species were assembled, providing excellent references for all omics analyses. Using the data from the experiments on the above mentioned research samples, we determined the attractive VOC and detected the genes related to the biosynthesis of this VOC, and identify the molecular mechanisms contributing to the host-specificity of the pollinator. More importantly, based on the omics data and function validation experiments, we found some molecular mechanisms underpinning coadaptation in fig-pollinator mutualisms.
Sampling strategy	Because prior knowledge showed that the attractive VOCs is only emitted by figs at receptive stage (see Fig. 1a), we therefore collected tissues (for transcriptome and proteome analysis) and VOCs from figs at prereceptive and receptive stages to detect the attractants as well as the related genes. Moreover, both VOC synthesis and chemical defense patterns may vary between the ostiolar tissues and the enclosed female florets (see Fig. 1a), and thus we sampled both types of fig tissues for comparative analysis. As to the pollinator, we first conducted behavior preference experiment to help identify the attractive VOCs, and then analyzed the gene families relevant to olfactory and detoxification and tested whether these genes expressed using transcriptome and proteome analysis. This comparison allowed us to explain why the pollinator larvae can live inside host figs and how adult pollinators detect receptive figs. We selected five individuals for each sex of Ficus pumila var. pumila and five replicates of host plants for Wiebesia pumilae (Supplementary Table 7) to satisfy the requirement of sample size for statistical analysis.
Data collection	VOC collection, fig tissue sampling for Ficus pumila var. pumila figs at both prereceptive and receptive stages and function validation of key genes were conducted by Yu Zhang, Jing-Jun Yang, Shan Chen, Rong Wang and Yang Yang. Sampling of individuals of Wiebesia pumilae was performed via collecting mature figs by Rong Wang, Yang Yang, Yuan-Yuan Ding and Xin Tong. Experiments of VOC component analysis, electrophysiological responses and behavioral preference of fig wasps were conducted by Yu Zhang, Gang Wang, Ping Wen and Chen Shan. Sequencing of genomes, transcriptomes and proteomes and metabolomic data were carried out using Illumina and Pacbio platforms as well as LC-MS-MS in BGI Genomics, and Qing-Feng Liu, Rong Wang, Yi Jing, Zhen Yue, Huan- Ming Yang and Ye Yin recorded the data.
Timing and spatial scale	Data collection in this study started from April 2017 and finished in September 2018. In April 2017, we selected the experimental plants and sampled fig tissues for transcriptome and proteome analysis, and we also collected leaves of the plant and adult females of the pollinator for genome assembly. From May 2017 to March 2018, high-quality genomes of both the plant and the pollinator were sequenced and assembled, and transcriptome and proteome analysis for fig tissues was conducted. In April 2018, we conducted VOC collection and fig tissue sampling. From May to September 2018, we performed VOC component analysis, electrophysiological responses and behavioral preference of fig wasps, and transcriptome and proteome analysis. From March to August 2020, we performed experiments to validate the functions of key genes and obtain metabolomic data to compare the chemical profiles between different types of tissues of figs. In case of large spatial variations in our studied traits, we only selected the experimental plants in a restricted area of c. 20 square kilometers.
Data exclusions	No data were excluded.
Reproducibility	Replications were set for all experiments, and all attempts at replication were successful.
Randomization	In this study, we chose the experimental plants and the mature figs haphazardly, from which fig wasps were sampled.

Bl	inc	ling
----	-----	------

Different groups of researchers and postgraduates conducted the experiments in each year, without being informed of the results obtained from other experiments in this study. Analyses were carried out using open source programmes, and therefore the algorithms were not specifically designed for our data.

Did the study involve field work? Xes No

Field work, collection and transport

Field conditions	The area for field work was mainly covered by subtropical evergreen broadleaved forests and was dominated by subtropical monsoon climate with the average temperature of 17 degrees Celsius and the mean precipitation of 1400 mm. As a climbing plant species, Ficus pumila var. pumila mainly grows on the trees at forest edges and on walls of old houses in villages. During the field work period, we collected VOCs from experimental plants for testing pollinator preference and sampled fresh plant tissues and W. pumilae individuals for transcriptome and proteome analysis.
Location	All field work was conducted in close proximity to Zhejiang Tiantong Forest Ecosystem National Observation and Research Station (E 121°47', N 29°48').
Access & import/export	The area for field works is not included in any nature reserves, and therefore no permits are required.
Disturbance	No serious disturbance was caused in this study because we only sampled some plant tissues and a few individual pollinating insects, without destroying any of the experimental plants and their crops.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

M	let	hoo	st
1.4		100	20

n/a	Involved in the study	n/a	Involved in the study
	Antibodies	\boxtimes	ChIP-seq
\boxtimes	Eukaryotic cell lines	\boxtimes	Flow cytometry
\boxtimes	Palaeontology and archaeology	\boxtimes	MRI-based neuroimaging
	Animals and other organisms		
\boxtimes	Human research participants		
\boxtimes	Clinical data		
\boxtimes	Dual use research of concern		

Antibodies

Antibodies used	To examine whether the key genes associated with the biosynthesis of the VOC attracting pollinating wasps are under the regulation of some transcription factors (HD-ZIP (FPUM_007795-RA), HD-ZIP (FPUM_003657-RA), bZIP (FPUM_016019-RA) and bHLH (FPUM_004914-RA)), we conducted ChIP-qPCR experiments. We therefore purchased polyclonal antibodies for immunoprecipitations from ABclonal Biotechnology (https://www.abclonal.com.cn/).
Validation	The function of each polyclonal antibody was tested using ChIP-qPCR experiments to show whether it can immunoprecipitate the predicted target transcription factor.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Our study did not involve laboratory animals, though four Japanese White Rabbits were used in producing polyclonal antibodies (that were purchased from ABclonal Biotechnology (https://www.abclonal.com.cn/)) for immunoprecipitation, an essential step in ChIP- qPCR.
Wild animals	For de novo assembly of Wiebesia pumilae genome, tests of behavior preference and transcriptome and proteome analysis, we collected adults of W. pumilae from the experimental Ficus pumila var. pumila plants in the vicinity of Zhejiang Tiantong Forest Ecosystem National Observation and Research Station in April 2017 and 2018 and from a F. pumila var. pumila plant located in South China Botanic Garden in April 2017 and 2020. Permission for sampling this insect species is not required because it is not an endangered species.
Field-collected samples	All sampled individuals of Wiebesia pumilae except those used in experiments of electrophysiological responses and behavior preference were immediately stored into liquid nitrogen. The pollinators used in the two laboratory experiments were raised within their natal figs at the conditions of c. 20 degrees Celsius and c. 70% relative humidity.
Ethics oversight	For conducting ChIP-qPCR experiments, we purchased polyclonal antibodies from ABclonal Biotechnology (https://

www.abclonal.com.cn/), which has the permission of using Japanese White Rabbits (permission number: SYXK 2017-0065) authorized by Department of Science and Technology of Hubei Province, China. No ethical approval or guidance was required for collecting tissue samples from F. pumila and W. pumilae.

Note that full information on the approval of the study protocol must also be provided in the manuscript.