







# Genomic evidence of prevalent hybridization throughout the evolutionary history of the fig-wasp pollination mutualism

Gang Wang <sup>1,15</sup>✉, Xingtian Zhang<sup>2,3,15</sup>, Edward Allen Herre<sup>4</sup>, Doyle McKey <sup>5</sup>, Carlos A. Machado<sup>6</sup>, Wen-Bin Yu <sup>7</sup>, Charles H. Cannon <sup>8</sup>, Michael L. Arnold<sup>9</sup>, Rodrigo A. S. Pereira<sup>10</sup>, Ray Ming <sup>11</sup>, Yi-Fei Liu <sup>12</sup>, Yibin Wang<sup>13</sup>, Dongna Ma<sup>14</sup> & Jin Chen<sup>1,3</sup>✉

*Ficus* (figs) and their agaonid wasp pollinators present an ecologically important mutualism that also provides a rich comparative system for studying functional co-diversification throughout its coevolutionary history (~75 million years). We obtained entire nuclear, mitochondrial, and chloroplast genomes for 15 species representing all major clades of *Ficus*. Multiple analyses of these genomic data suggest that hybridization events have occurred throughout *Ficus* evolutionary history. Furthermore, cophylogenetic reconciliation analyses detect significant incongruence among all nuclear, chloroplast, and mitochondrial-based phylogenies, none of which correspond with any published phylogenies of the associated pollinator wasps. These findings are most consistent with frequent host-switching by the pollinators, leading to fig hybridization, even between distantly related clades. Here, we suggest that these pollinator host-switches and fig hybridization events are a dominant feature of fig/wasp coevolutionary history, and by generating novel genomic combinations in the figs have likely contributed to the remarkable diversity exhibited by this mutualism.

<sup>1</sup>CAS Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, Yunnan, China. <sup>2</sup>Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China. <sup>3</sup>Center for Plant Ecology, Core Botanical Gardens, Chinese Academy of Sciences, Mengla, Yunnan, China. <sup>4</sup>Smithsonian Tropical Research Institute, Balboa, Ancon, Republic of Panama. <sup>5</sup>CEFE, University of Montpellier, CNRS, University Paul Valéry Montpellier 3, EPHE, IRD, Montpellier, France. <sup>6</sup>Department of Biology, University of Maryland, College Park, MD, USA. <sup>7</sup>Center for Integrative Conservation, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, Yunnan, China. <sup>8</sup>Center for Tree Science, The Morton Arboretum, Lisle, IL, USA. <sup>9</sup>Department of Genetics, University of Georgia, Athens, GA, USA. <sup>10</sup>Department of Biology, FFCLRP, University of São Paulo, Ribeirão Preto, São Paulo, Brazil. <sup>11</sup>Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA. <sup>12</sup>College of Pharmacy, Hubei University of Chinese Medicine, Wuhan, Hubei, China. <sup>13</sup>Center for Genomics and Biotechnology, Fujian Agriculture and Forestry University, Fuzhou, Fujian, China. <sup>14</sup>Key Laboratory of the Ministry of Education for Coastal and Wetland Ecosystems, College of the Environment and Ecology, Xiamen University, Xiamen, Fujian, China. <sup>15</sup>These authors contributed equally: Gang Wang, Xingtian Zhang. ✉email: [wanggang@xtbg.org.cn](mailto:wanggang@xtbg.org.cn); [cj@xtbg.org.cn](mailto:cj@xtbg.org.cn)

Hybridization is increasingly recognized as making important contributions to adaptation and diversification<sup>1–4</sup>. For example, polyploid hybrid diversification is widely accepted in plants, as polyploidy results in rapid reproductive isolation from parental lineages<sup>4,5</sup>. However, the degree to which homoploid hybridization and adaptive introgression contribute to diversification is still actively debated, owing in part to the lack of direct, identifiable reproductive isolation mechanisms<sup>4,6,7</sup>. Animal pollination also is thought to contribute to the diversification of flowering plants, primarily by facilitating reproductive isolation via diversifying pollinator niches and pollinator selection<sup>8–12</sup>. Additionally, animal pollination also facilitates (homoploid) hybrid diversification of flowering plants by generating pollinator-mediated hybrids, followed by the subsequent establishment of pollinator-mediated isolation<sup>6,13–15</sup>. However, empirical studies that focus on the potential for interaction between animal pollination and hybridization to affect and contribute to evolutionary diversification are rare.

Obligate pollination systems, in which plants of a specific taxonomic group are exclusively pollinated by a specific taxonomic group of pollinators, provide an opportunity for detailed examination of these evolutionary interactions<sup>16–20</sup>. These extreme obligate-pollination plants are often more diverse than their non-obligate sister groups<sup>16,18,20,21</sup>. It has been generally accepted that tight pollinator specificity has accelerated diversification, primarily through a mechanism of enhanced reproductive isolation<sup>9,16–18</sup>. However, over the last few decades, an increasing number of empirical studies have demonstrated that, even in extreme obligate pollination systems, pollinator sharing, host switches, and interspecific hybridization occur<sup>14–16,22</sup>, suggesting that strict one-to-one coevolution is not the only evolutionary process occurring in these systems. Thus, a more diffuse model involving pollinator switches and interspecific hybridization is potentially more appropriate for describing speciation and diversification in these “strict-sense” obligate pollination systems<sup>14,15</sup>.

The genus *Ficus* (figs, Moraceae) is one of the most diverse woody plant genera (~800 described species)<sup>21</sup>. With their pollinating fig wasps (Hymenoptera: Chalcidoidea: Agaonidae), they represent perhaps the most extreme and ancient (~75 million years) obligate pollination mutualisms known<sup>20,23</sup>. Figs are distributed in tropical and subtropical regions around the world, with the highest diversity of subgenera and species occurring in South-East Asia<sup>21,24</sup>. Given the observation that morphologically defined taxa (subgenera and sections) of figs are usually pollinated by distinct genera of pollinators, early speculation on the mode of fig and wasp co-evolution emphasized the role of extreme host-specificity of the pollinator wasps coupled with strict-sense coevolution between individual species of figs and their pollinators<sup>17,20,25,26</sup>.

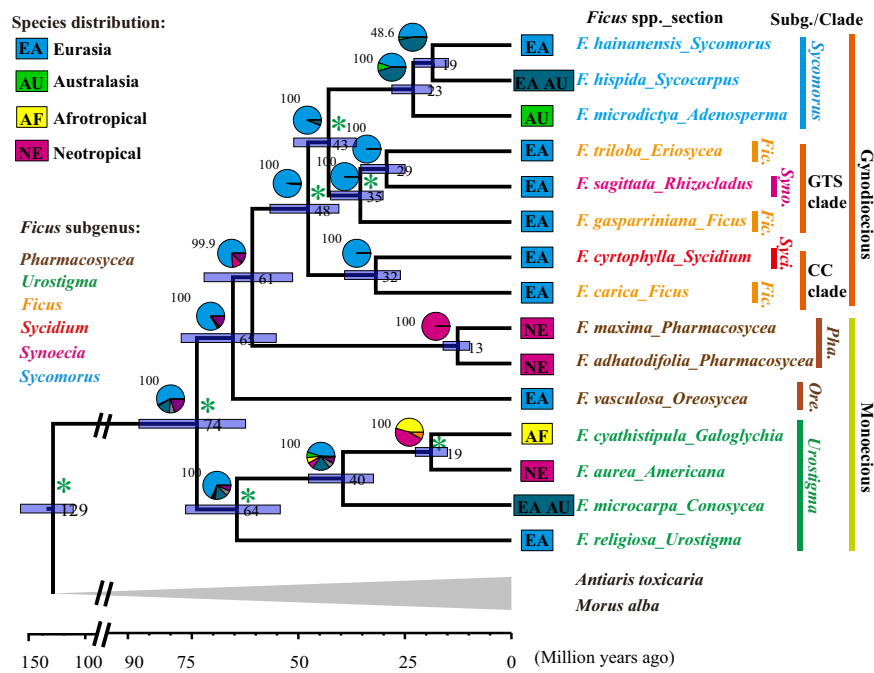
However, broader species sampling combined with increasingly detailed and comprehensive genetic analyses of both host and pollinator species have identified ever larger numbers of exceptions to the previously assumed extreme host specificity. Cases in which multiple wasp species are associated with a single host, as well as cases in which wasp species are shared among two or more hosts, have not only been discovered but also appear to be relatively common<sup>27–31</sup>. Further, in studies of sympatric, closely related figs (within a section) and their respective pollinator species (within a genus)<sup>32,33</sup>, clear cases of phylogenetic incongruence have been documented and are consistent with host-switching that can lead to the apparent hybridization and introgression found among some closely related fig species<sup>32,34–37</sup>. Host switching leading to hybridization is also suggested by studies of distantly related fig and wasp taxa (across sections of figs and genera of pollinators). Specifically, apparent

incongruence of both the fig–fig wasp phylogenetic relationships and the nuclear–chloroplast associations of figs across sections of the genus is consistent with the possibility of hybridization at even these deeper taxonomic scales<sup>15,20,32,38</sup>. Collectively, the available data suggest that a more diffuse model of coevolution that incorporates frequent host switching followed by hybridization and introgression might better explain the evolutionary dynamics in this system than strict-sense co-speciation<sup>15,22,33,34</sup>. However, the phylogenetic inferences of much of the existing (suggestive) work are based on data from relatively few, marginally informative genes that only weakly support crucial basal nodes. Further, these studies lack both genetic depth (e.g., entire nuclear and organelle genomes) and a comprehensive sampling design (e.g., taxa that represent what are thought to be the major evolutionary transitions in the genus *Ficus*). Finally, rigorous tests to distinguish true genetic introgression from the alternative process of incomplete lineage sorting (ILS) have generally been lacking.

In this work, we collect complete sequences for nuclear, chloroplast, and mitochondrial genomes from 15 fig species representing all major recognized *Ficus* clades. Previous studies suggest that these clades have been diverging from each other for roughly 75 million years<sup>20</sup>. Integrating complete sequences for two other Moraceae species used as outgroups, we apply multiple analytical approaches to reconstruct phylogenetic relationships among these clades and identify cases of hybridization potentially recorded in the nuclear genomes throughout the evolutionary history of figs. We then compare the inferred evolutionary histories of the nuclear, chloroplast, and mitochondrial genomes across these 15 fig species. Finally, we use published phylogenies of the corresponding pollinator species<sup>20,23</sup> to assess the evidence for pollinator switches throughout the course of *Ficus* and pollinator wasp evolution. Here, we find direct evidence indicating frequent hybridization and introgression events in fig nuclear genomes. These events are detected even among distantly related, currently non-sympatric lineages of figs, suggesting a process that has been acting throughout the evolutionary history of figs. Specific inferred hybridization events are corroborated by multiple analytical approaches (ABBA-BABA *D*-statistics, Phylo-Networks, and Bayesian Concordance). Furthermore, this inference is also supported by the significant lack of concordance of the topologies of the chloroplast and mitochondrial phylogenies of the 15 *Ficus* species with any supported nuclear phylogenies, or, indeed, with each other. Finally, we show that no currently supported published wasp phylogeny corresponds to any fig phylogeny based on nuclear, chloroplast, or mitochondrial data. Significantly, most hybridization events detected among *Ficus* main clades exhibit associated pollinator host-switch events. We conclude that the evolutionary history of figs and wasps has likely been characterized by frequent host switches of the pollinators and that this process has regularly produced novel genetic combinations within and among host fig genomes.

## Results

**Sequencing and reconstruction of *Ficus* evolution.** We used the 13 chromosomes of the *Ficus microcarpa* genome as a reference template<sup>39</sup> (Supplementary Note 1) to identify the nuclear single nucleotide polymorphisms (SNPs) from 17 Moraceae samples, including 15 fig species representing all main *Ficus* clades (Supplementary Tables 1 and 2). We identified 1,813,332 SNPs from 24,188 genes (including 848,552 SNPs in exons and 964,780 in introns) and 2,660,312 SNPs from intergenic regions, from a total of 196.7 Gbp of raw sequencing data. The average sequencing depth was 26.91× per sample, and 72.76× per SNP locus. The sequenced reads of each *Ficus* sample covered most of the



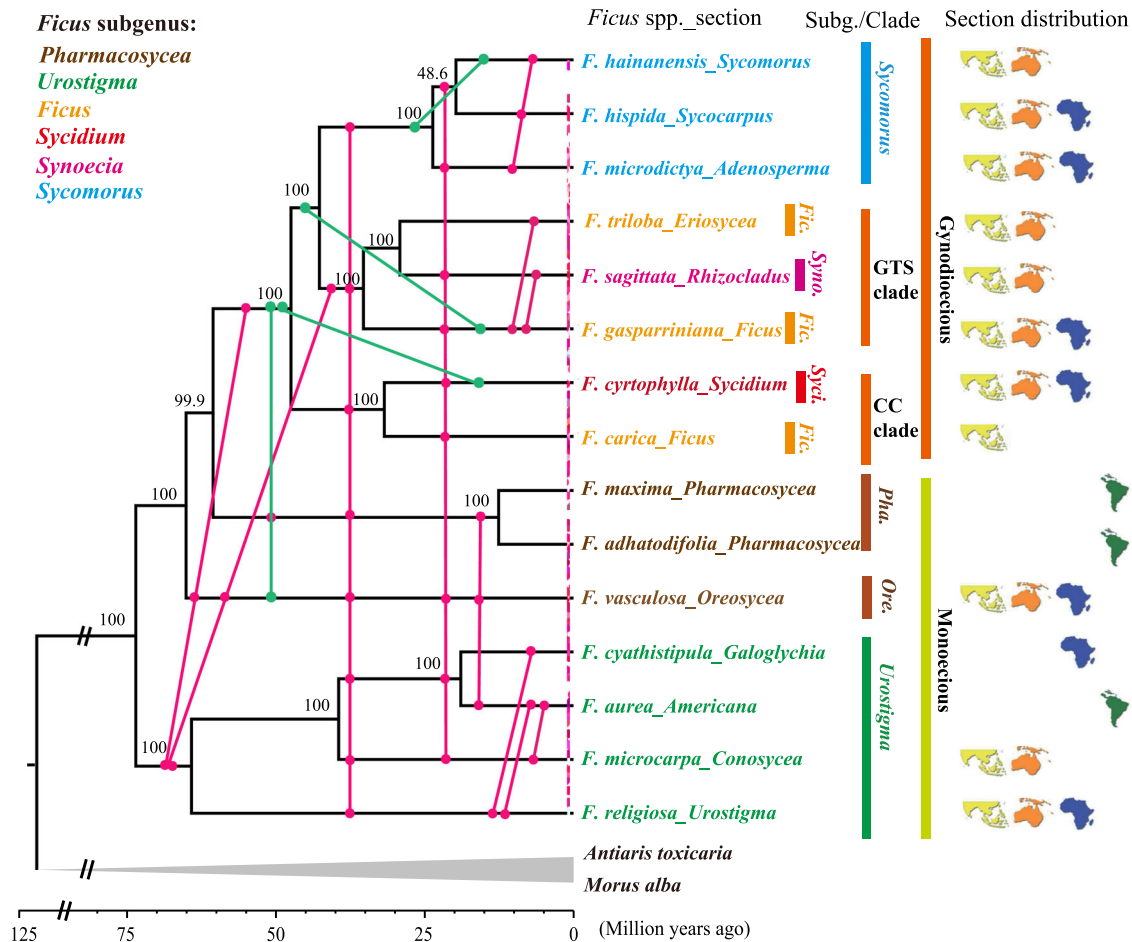
**Fig. 1 *Ficus* Astral species tree (based exclusively on nuclear data) with inferred divergence times and ancestral geographic ranges of *Ficus*.** The Astral *Ficus* species tree was inferred based on non-overlapping 100-kb genomic windows, with almost all nodes being highly supported with bootstrap values >99, excepting one (48.6). The bootstrap values are indicated above the pie chart at each node. Values and bars on the nodes indicate the mean and 95% credibility intervals of node age. The pie charts indicate the relative probability of possible geographic ranges estimated. Current distribution range of related section (tip symbols) and the subgenus information (tip name color) of sampled species are also shown. Green asterisks indicate calibrated nodes used for inferring divergence times. Ranges that are combinations of these four distributions are indicated by colors formed by mixing of the four component distribution colors. Both species distribution and taxonomic classification follow Cruaud et al.<sup>20</sup>. *Ficus* is inferred to have originated in Eurasia and all main clades have more than 10 million years history of coexistence in Eurasia, before dispersal to other areas. This ancestral sympatry could have permitted ancient hybridization among them. Refer to Supplementary Figs. 1–3, Supplementary Table 3, and Appendices 1, 3, and 4 for details.

reference genome (*F. microcarpa*), ranging from 88% to 99%. We assembled the circular genomes of both the mitochondria and chloroplasts from the same set of samples for all species. Chloroplast genome size in *Ficus* is consistently around 160 kb (kilobase), while that of mitochondria varies from 430 kb to 711 kb (Supplementary Table 2).

Using the nuclear genomic level dataset, coalescent-based Astral species trees<sup>40</sup> were inferred for *Ficus* based on four types of nuclear genomic sliding windows generated with partition strategies of non-overlapping windows of size 500, 100, and 50 kb, and a SNPs-fixed window of 1000 SNPs (see “Methods” and Supplementary Table 3 for details). The four Astral species trees corresponding to each window size generated similar *Ficus* topologies with high bootstrap values at almost all nodes (Fig. 1, Supplementary Fig. 1A–D, and Appendix 1). All four Astral species trees supported the monophyletic subgenus *Urostigma* (monoecious, pantropical, with both active and passive pollination, Supplementary Note 2) as sister to all other *Ficus* species. After this basal split, the next split supports the two sections of subgenus *Pharmacosycea* (monoecious, with passive pollination for section *Pharmacosycea* and active for section *Oreosycea*) appearing as sister taxa to the monophyletic gynodioecious figs, within which the monophyletic subgenus *Sycomorus* is nested. Similar to results of all previous studies<sup>20,38,39,41</sup>, the two sections of subgenus *Pharmacosycea* did not constitute a monophyletic group. Within the gynodioecious clade, subgenus *Ficus* (part of the gynodioecious clade) also emerged as polyphyletic, while subgenus *Sycomorus*, subgenera *Ficus* + *Synoecia* group (“GTS-clade” in Figs. 1 and 2) and the subgenus *Sygidium* + *F. carica* group (“CC-clade” in Figs. 1 and 2) were recovered as monophyletic with strong support (Fig. 1 and Supplementary Fig. 1).

Consistent with a previous study that included 200 species<sup>20</sup>, our dating and ancestral area reconstructions using the Astral tree based on the complete nuclear genomic data suggest that *Ficus* originated in Eurasia roughly 73.83 (62.39–87.41) million years ago (Mya) (Supplementary Figs. 2A, 3A, Supplementary Table 4, and Appendices 3, 4). The Astral tree-based reconstructions further suggest that all main fig clades coexisted in Eurasia for more than 10 million years. This period of sympatric coexistence appears to have been followed by three major events of expansion out of Eurasia: (1) at ~60.81 (51.35–72.12) Mya, section *Pharmacosycea* colonized the Neotropics, (2) at ~39.52 (32.38–47.60) Mya, the common ancestor of sections *Americana* and *Galoglychia* colonized the Neotropics and Afrotropics, respectively, and (3) at ~23.05 (18.83–28.06) Mya, section *Adenosperma* colonized Australasia (Supplementary Fig. 3).

However, using the same nuclear genome dataset, analyses using a Bayesian concordance approach (BCA)<sup>42</sup> suggest an alternative topology for the deepest node of *Ficus* (Supplementary Fig. 1E and Appendix 2). The primary concordance tree (PCT tree, inferred with the program BUCKY<sup>43</sup>) instead suggests section *Pharmacosycea* is best supported as sister to all other figs<sup>39</sup> (concordance factor (CF) = 0.406, meaning that 40.6% of genomic windows support this topology<sup>44</sup>), compared to any other single clade (Supplementary Table 5). Apart from the placement of the basal section, the remaining topological structure of the PCT tree is similar to those of the Astral trees (Supplementary Fig. 1). Divergence times and ancient biogeography inferred based on the PCT tree suggest that *Ficus* originated in either Eurasia or the Neotropics roughly 63.20 (58.70–71.15) Mya (Supplementary Figs. 2B, 3B and Appendices 3, 4), followed by a rapid (~1 My) divergence between the



**Fig. 2 Hybridization events within *Ficus* inferred with ABBA-BABA *D*-statistics and PhyloNetworks exhibited on the Astral tree.** The Astral *Ficus* species tree (indicated with horizontal lines) inferred based on the non-overlapping 100-kb-windows nuclear genome datasets, with bootstrap values at nodes. The vertical lines in the tree represent the hybridization events inferred with ABBA-BABA *D*-statistics alone (rose-red lines and dots) or with both the *D*-statistics and PhyloNetworks (green lines and dots). For all pairs of species or clades with a dot where a vertical line crosses the (horizontal) branch indicating the species or clade, hybridization was detected involving taxa connected by the same vertical line. The dashed rose-red line near the tips of branches indicates that signals of hybridization could be detected among almost all species across different subgenera with *D*-statistics. Points where vertical lines contact branches do not reflect the precise estimated time of hybridization, but rather a time interval (between stem and crown age of related taxa) during which hybridization occurred. Current distribution range of related section (map symbols) and the subgenera information (tip name color) of sampled species are also shown. *Ficus* classification and current distribution at the section level are also shown. Refer to Supplementary Data 1 and 2 and Appendices 2 and 6 for details.

Neotropical section *Pharmacosycea* and all other figs (which dispersed to or originated in Eurasia). Machado et al.<sup>23</sup> found support for a similar scenario for pollinating wasps.

Maximum likelihood (ML) trees of both chloroplast and mitochondrial genomes based on full genome sequences were highly supported at almost all nodes. However, both showed striking and significant differences from the nuclear phylogenies (Fig. 2, Supplementary Fig. 4, and Appendix 5), as well as from each other. Consistent with the PCT tree based on nuclear genomes, and with most previous *Ficus* phylogenies<sup>20,26,38,39,41</sup>, both organelle-based phylogenies strongly supported section *Pharmacosycea* (monoecious, Neotropical, passive pollination) as being sister to all other *Ficus*. Further, none of the six currently recognized subgenera of *Ficus* (subgenera *Urostigma*, *Pharmacosycea*, *Ficus*, *Sycidium*, *Synoecia*, and *Sycomorus*)<sup>21</sup> was supported as a monophyletic clade by either of the two organelle-based phylogenies<sup>38</sup>.

Analyses based on the Astral tree and PCT tree both support the interpretation that all main clades of *Ficus* co-occurred sympatrically during portions of their biogeographical histories.

However, Astral and PCT trees suggest different clades as being the sister group to the rest of *Ficus* (subgenus *Urostigma* and section *Pharmacosycea*, respectively). Given its strong agreement with PhyloNetworks<sup>45</sup> analyses (see below) (Supplementary Fig. 6), we will primarily focus on the inferences based on the Astral tree topology for the remaining analyses. However, similar analyses were also conducted based on the alternative PCT tree topology and yielded similar results that suggest prevalent hybridization events.

**Detecting hybridization and distinguishing it from ILS.** In order to detect potential hybridization events, and distinguish these events from other processes such as ILS that potentially produce similar genetic signatures<sup>44–48</sup>, we first conducted ABBA-BABA *D*-statistics analysis<sup>46,47</sup>. We conducted these analyses for four-taxon groups hierarchically, from the species tips (fine scale) to the main clades (coarse scale) of the *Ficus* Astral species tree. These analyses yielded significant putative hybridization signals in most four-taxon groups across different

phylogenetic/taxonomic levels (e.g., *Ficus* main clades, subgenus, sections, or species level) (Fig. 2 and Supplementary Data 1). All surveyed species, sections, subgenera, or main clades exhibited at least one hybridization event with other clades. Most inferred hybridization events occurred among groups that currently exhibit overlapping geographical distributions spanning Eurasia to Australasia. However, at least a few hybridization events were also identified between groups currently found on different continents: (i) Eurasian and Australasian sections and each of the exclusively New World sections *Pharmacosycea* and *Americana*; (ii) Eurasian and Australasian sections and the exclusively Afrotropical section *Galoglychia*; and (iii) *Galoglychia* and each of the New World sections (Fig. 2).

A second analytical approach, PhyloNetworks<sup>45</sup>, detected four hybridization events among *Ficus* clades (Supplementary Fig. 6 and Appendix 6). All four events were also supported by the ABBA-BABA *D*-statistics (Fig. 2, green lines). Notably, one ancient hybridization event was detected that suggested nearly 41% of the *Oreosycea* (monoecious) genome was transferred to the ancestor of the gynodioecious clade. The largest introgressive genomic portion of the gynodioecious clade detected with PhyloNetworks was also supported by the largest absolute *Z*-value (346.15) detected in ABBA-BABA *D*-statistics, from the hybridization event between the gynodioecious clade and the *Oreosycea* clade (Supplementary Data 1).

Similarly, Bayesian concordance analysis results with BUCKy<sup>44,48</sup> also inferred frequent hybridization events (Supplementary Fig. 7 and Supplementary Data 2). Consistent with the other two methods, many of the hybridization events BUCKy detected appear to have occurred between clades from different subgenera, or even between different sexual systems (monoecious and gynodioecious) (Supplementary Fig. 7, Supplementary Data 2, and Appendix 2). We note that all 12 hybridization events detected by BUCKy were corroborated by results of ABBA-BABA *D*-statistics analyses (Supplementary Fig. 7 and Supplementary Data 1). In the BUCKy analyses, all *Ficus* main clades were inferred to be involved in at least one hybridization event, with the exception of section *Pharmacosycea*. We note that section *Pharmacosycea* (which the PCT tree suggests it sister to all other figs) currently exhibits an exclusively New World distribution.

**Phylogenetic incongruence among the three genomic compartments of *Ficus*.** Regardless of whether the Astral or PCT topology was used, we repeatedly detected significant phylogenetic incongruence among the nuclear, chloroplast, and mitochondrial genomes of *Ficus*. Therefore, the hypothesis of strict codivergence (i.e. that pairs of genomic compartments diverge at the same time and are continually associated) among them was clearly rejected (Fig. 3a–c). We then ran event-based cophylogenetic analysis using JANE 4.0 (ref. 49) to infer the relative contributions of different potential evolutionary events in generating phylogenetic incongruence (e.g., duplication, associate switches, associate losses, and failure to diverge). Notably, JANE 4.0 inferred high frequencies of both associate switch (suggestive of hybridization<sup>50</sup>, as organellar genome lineages switch among different background nuclear genome lineages) and associate loss events (“parasite lineage” [e.g., organellar genome] is associated with only one of two divergent “host lineages” [e.g., nuclear genomes], while another “host lineage” was associated with a different “parasite lineage”) (Fig. 4a, Supplementary Tables 6, 7, and Appendix 7). Furthermore, detailed analyses of the nuclear and the two organelle-based phylogenies uncovered frequent, relatively independent switching of both organelle genomes across all *Ficus* subgenera (Fig. 5a and Supplementary Figs. 8–10). This finding is consistent with previous demonstrations of

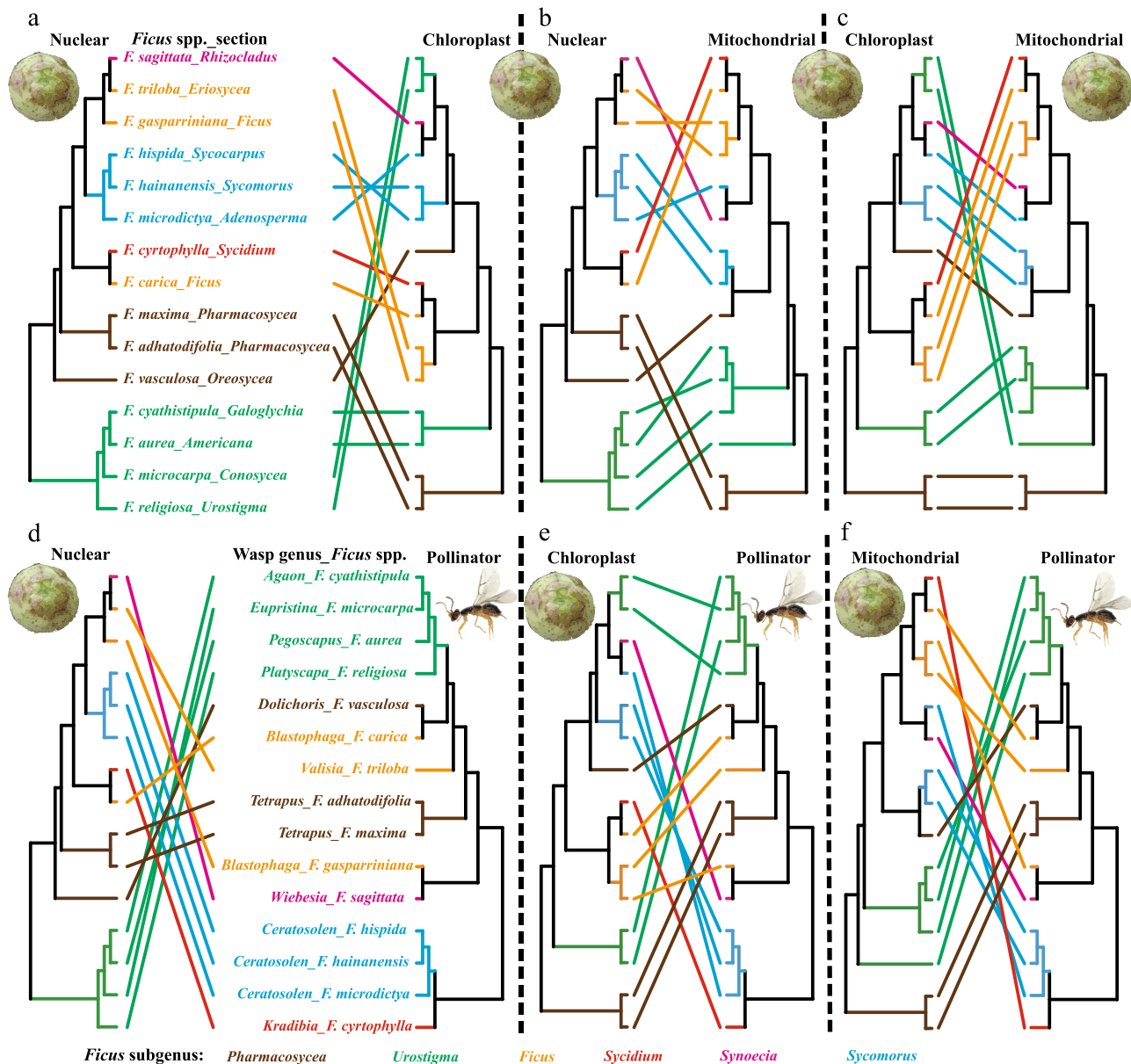
biparental inheritance and heteroplasmy in plant organelles<sup>50–52</sup>. Finally, our analyses of the geographic distribution of each of the 59 fig species (Supplementary Data 3) included in Bruun-Lund et al.’s chloroplast genomic phylogeny<sup>38</sup> showed an intriguing pattern. The chloroplast genomes of what appeared to be distantly related fig species from clearly different subgenera were frequently clustered together in the chloroplast genomic phylogeny. This unexpected pattern is particularly true for taxa that currently share the same geographic region. All of these findings are consistent with genetic exchange of organelles among lineages more likely to have experienced some level of sympatry (Supplementary Fig. 11).

**Pollinator host-switches and the evolutionary co-divergence of *Ficus* and fig wasps.** Analogously to the incongruence found among fig nuclear and both organelle-based phylogenies, clear phylogenetic incongruences were also detected between the host figs and their associated pollinators (Fig. 3d–f). We conducted co-phylogenetic analyses using all combinations of the best supported Astral tree, the PCT tree and the two organellar phylogenies generated here for *Ficus* and the two most widely accepted (but substantially different) published fig wasp phylogenies<sup>20,23</sup> (Supplementary Fig. 5 and Supplementary Note 3). Regardless of which of the two wasp phylogenies was tested against the various nuclear and organelle-based *Ficus* phylogenies, pollinator host-switch events were always inferred at higher frequency than co-speciation events (Fig. 4b–c, Supplementary Table 7, and Appendix 7). Further, reconciliation between any *Ficus* phylogeny and any pollinator phylogeny revealed that almost all *Ficus* main clades have been repeatedly involved in pollinator host-switching events (Fig. 5b and Supplementary Figs. 12–15). Meanwhile, almost all of the inferred pollinator host-switching events correspond to the hybridization events identified above (Fig. 5b, Supplementary Table 7, and Supplementary Note 4), and vice versa, most hybridization events among main clades (18 of 22 listed in Supplementary Table 8) also corresponded to one or more pollinator host-switching events.

## Discussion

The complete nuclear, mitochondrial, and chloroplast genomes for 15 species of figs (*Ficus* spp.) that we collected represent the most comprehensive genetic sampling thus far conducted for key taxa of the plant hosts in what is arguably the most extreme case of obligate pollinator mutualism known. Despite the depth of genetic sampling, analyses of these data do not unambiguously resolve the phylogenetic relationships among the recognized *Ficus* clades, especially at important basal nodes. What these complete genomic data do unambiguously indicate is that hybridization and introgression of components of the nuclear genome, and exchange of entire mitochondrial and chloroplast genomes, have likely characterized *Ficus* throughout its evolutionary history. Moreover, no published phylogeny of the corresponding pollinator wasp species (Hymenoptera: Chalcidoidea: Agaonidae) corresponds to any strongly supported nuclear or organelle-based phylogeny of their hosts. Frequent pollinator host-switching, rather than strict-sense cospeciation, was inferred as the predominant event along fig–fig wasp coevolution. Collectively, the existing data suggest evolutionarily frequent host switches of the pollinators that introduce new combinations of both nuclear and organelle genes. This process appears to take place even among distantly related hosts and throughout fig–pollinator wasp coevolutionary history (Fig. 2 and Supplementary Table 8).

Multiple evidences support prevalent hybridization has occurred throughout the evolutionary history of *Ficus*. Frequent hybridization events were inferred to have taken place within and across all six *Ficus* subgenera, based on all three different types of

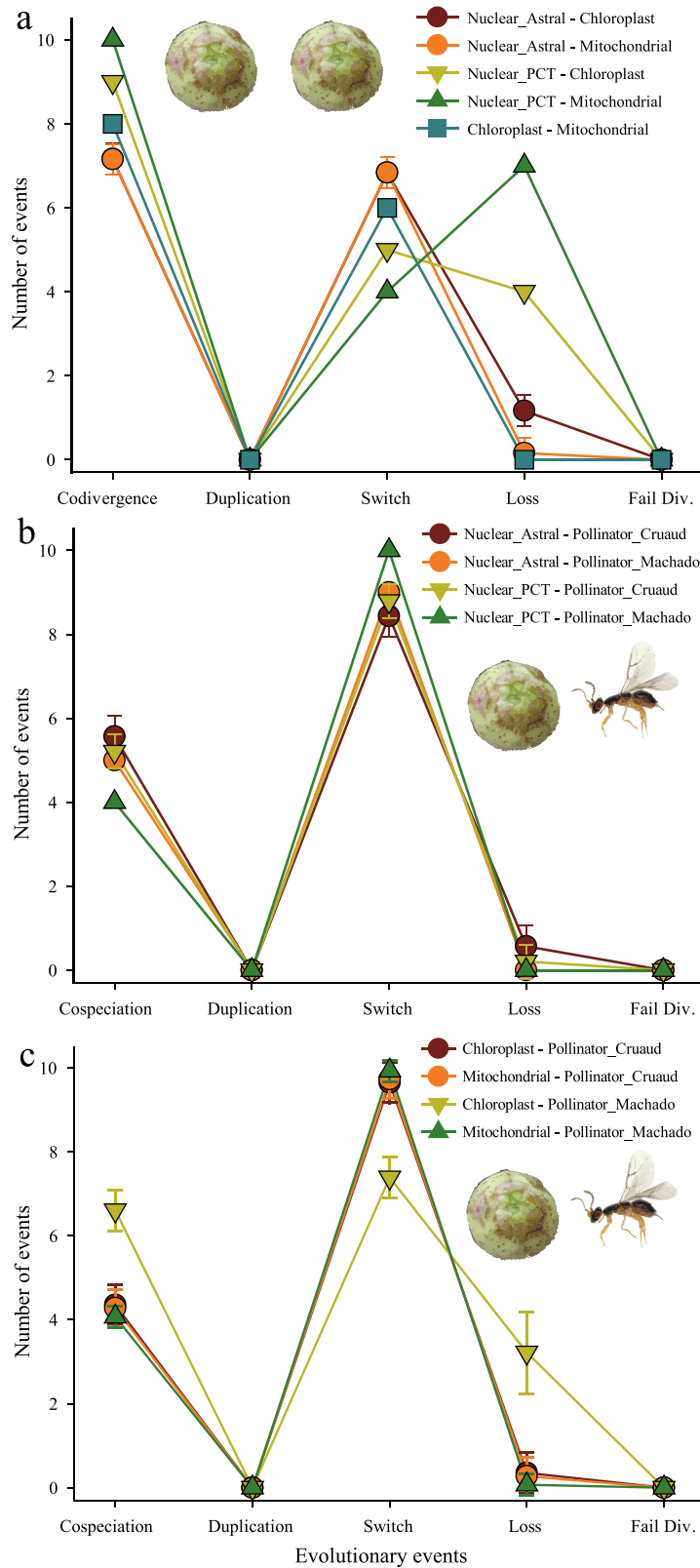


**Fig. 3 Phylogenetic incongruence among fig genomic compartments and pollinator.** Cophylogenetic comparison of each pair of phylogenies of three genomic compartments (a–c) and of the pollinators associated with the 15 fig species studied (d–f). *Ficus* and pollinator phylogenies used were indicated with symbol of figs and wasp separately. Tip-name colors indicate subgeneric classification. The pollinator phylogeny extracted from Cruaud et al.<sup>20</sup> was used here. Extensive phylogenetic incongruence is shown for each pair of phylogenies. Refer to Supplementary Figs. 1, 4, and 5a and Appendices 1 and 5 for details on the phylogenies used here.

analyses applied to nuclear genome data: ABBA-BABA *D*-statistics<sup>46,47</sup>, PhyloNetworks<sup>45</sup>, and BUCKY<sup>43,44</sup> (Fig. 2, Supplementary Fig. 7, and Supplementary Table 8). In particular, ABBA-BABA *D*-statistics<sup>46–48</sup> suggest extensive hybridization events in all 15 sampled species (14 sections) (Supplementary Data 1). Moreover, all hybridization events inferred with PhyloNetworks and BUCKY were also corroborated by results of ABBA-BABA *D*-statistics (Fig. 2 and Supplementary Table 8). Interestingly, several hybridization events suggest the possibility of ancient “ghost” hybridization events involving an extant taxon and a now-extinct clade<sup>47</sup> (Fig. 2 and Supplementary Figs. 6 and 7).

Further, neither the mitochondrial nor the chloroplast phylogenies are congruent with the nuclear species trees inferred by either ASTRAL or Bayesian approaches (BUCKY). They are not even congruent with each other (Figs. 3–5 and Supplementary Figs. 1 and 4). Both *Ficus* organelle genomes have clearly been

switching nuclear backgrounds throughout their history (Fig. 4 and Supplementary Table 6), consistent with the results of Bruun-Lund et al.<sup>38</sup>, and in ways that suggest that hybridization can result in both maternal and paternal inheritance of the organellar genomes in *Ficus*. We note that biparental and paternal inheritance of organellar genomes, as well as organelle heteroplasmy, have been shown in an increasing number of angiosperms<sup>50–52</sup>. In addition, our analyses of the published chloroplast genomic phylogeny of 59 fig species<sup>38</sup> revealed a geographic clustering pattern that often overrode expected similarity based on what subgenus the figs belonged to (Supplementary Fig. 11). This pattern suggests that hybridization among co-occurring species, even if they are otherwise distantly related, is the most likely mechanism accounting for the significant cyto-nuclear incongruence demonstrated in previous studies<sup>32,38</sup>. The simplest explanation consistent with all the above findings is that

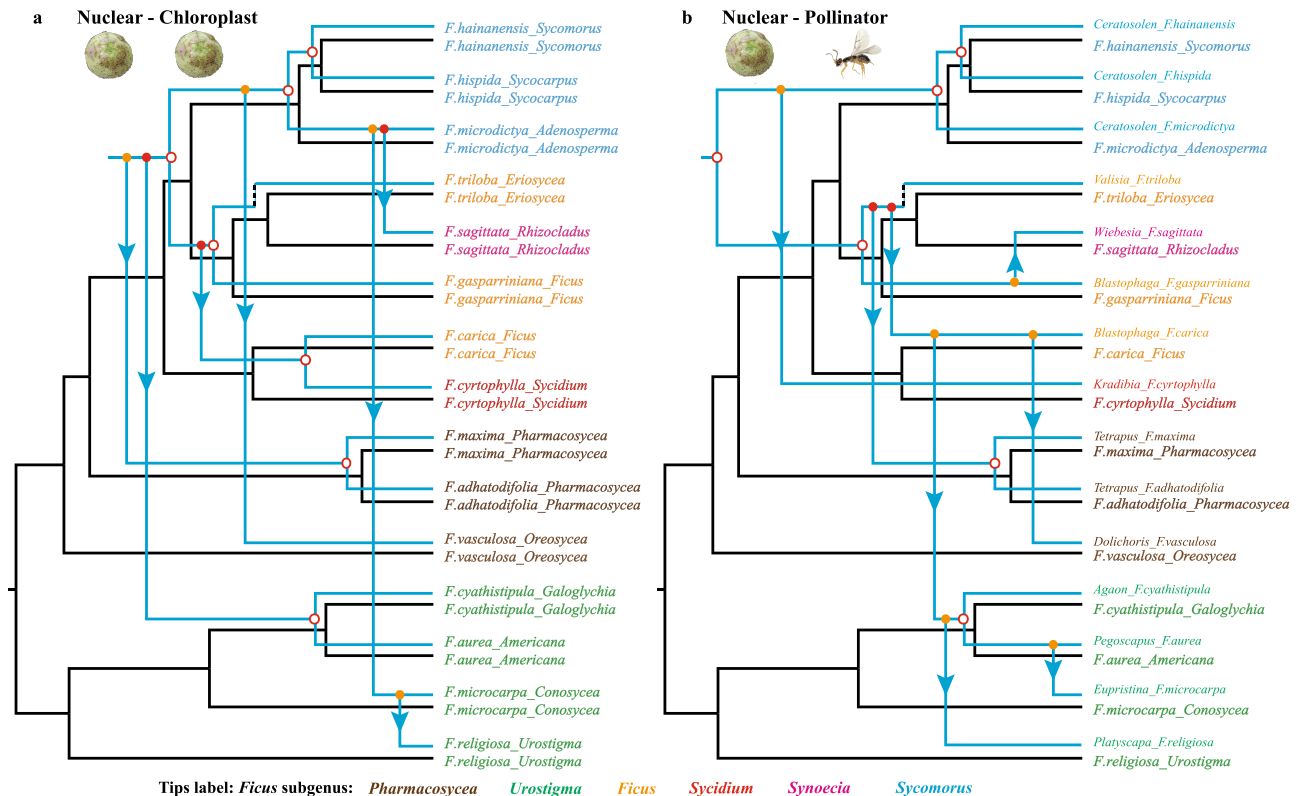


hybridization has occurred routinely throughout the evolutionary history of *Ficus*, and that its effects have been prevalent.

It is difficult to reconcile the complete genomic data for these 15 fig species with a standard biological species concept<sup>53</sup>, which assumes that species represent distinct gene pools through time, once speciation has occurred. In contrast, *Ficus* species may better fit the genic view of species, in which “species” exchange

potentially large regions of the genome, except possibly those regions controlling adaptive traits important for reproductive fitness<sup>54</sup>. Meanwhile, our study also provides evidence of hybridization at the deepest nodes (50–80 Mya) of a species-rich group. Such demonstrations are rare, perhaps because of what were previously methodological limitations<sup>55,56</sup>. The *Ficus* hybridization events inferred here also help reconcile the previously

**Fig. 4 Frequency of co-phylogenetic evolutionary events inferred by JANE.** Cophylogenetic analyses were conducted among phylogenies of the three genomic compartments (nuclear, chloroplast, and mitochondria) of *Ficus* (a) and then between these and genus-level phylogenies of the associated pollinators (b, c). *Ficus* and pollinator phylogenies used were indicated with symbol of figs and wasp separately. JANE takes the first tree as the “host tree”, and the second as the “parasite tree” shown in legends for cophylogenetic analysis. Means and standard errors were calculated with outcomes of all sets of solutions inferred with JANE under the same optimal cost. Lines linking numbers of different kinds of events are included for easy visualization; they add no quantitative information. Two *Ficus* nuclear phylogenies (Astral tree and primary concordance tree (PCT) in BUCKy) and two pollinator phylogenies<sup>20,23</sup> were used. Five kinds of events were inferred: codivergence/cospeciation, duplication, associate switch (genomic switches or pollinator switches), associate loss (organellar genome loss or pollinator extinction), and failure to diverge. Refer to Source Data 2, Supplementary Figs. 5, 8–10, 12–15, Supplementary Table 7, and Appendix 7 for details.



**Fig. 5 Co-phylogenetic reconciliations and detail evolutionary events inferred among main fig clades.** One of the equally most parsimonious co-phylogenetic reconciliations inferred with JANE under optimal cost setting for each of two kinds of phylogeny pairs: between *Ficus* phylogenies based on nuclear and chloroplast genomes (a) and between *Ficus* nuclear-genome phylogeny and the phylogeny of associated pollinators (b). *Ficus* and pollinator phylogenies used were indicated with symbol of figs and wasp separately. The black tree represents the nuclear genomic phylogeny in both a and b, and the blue trees represent chloroplast phylogeny (a) or pollinator phylogeny extracted from Cruaud et al.<sup>20</sup> (b). Tip labels indicate fig species and section in *Ficus* phylogeny, wasp genus, and associated host species in the pollinator phylogeny. Subgenus information (tip name color) of sampled species are also shown. Hollow red circles represent codivergence events in a and cospeciation events in b; solid colored circles and arrow lines represent events of associate switch (chloroplast switch) in a and of pollinator switch in b, while filled orange dots indicate existence of another potential switch location in the phylogeny that has equal cost, and a filled red dot means that all other potential switch locations have higher cost; dashed lines represent associate events of chloroplast loss (in a) or of pollinator extinction (in b). Refer to Supplementary Table 8 and Supplementary Fig. 5a for details.

encountered inconsistent phylogenetic reconstructions of *Ficus* that have been based on different morphological or molecular datasets, and analyzed with different methods that nonetheless assume separate gene pools<sup>20,57</sup>. Accordingly, this study highlights the utility of using multiple genomic compartments and multiple alternative analytical methods to distinguish between alternative phylogenetic hypotheses, and to corroborate inferences.

Pollinator host-switches could be a direct mechanism underlying *Ficus* hybridization events. Currently there are two main published and widely accepted molecular-based phylogenies of fig pollinating wasps<sup>20,23</sup> that correspond to our sampled host fig taxa. Although neither of these phylogenies is based on data with comparable depth to the whole genome data generated here for

key *Ficus* taxa, they represent the two currently most plausible, if alternative, views of fig wasp evolutionary history (Supplementary Fig. 5 and Supplementary Note 3). In each case, the data for both partners are sufficiently resolved to reject the hypothesis that either of the wasp phylogenies corresponds to any of the fig nuclear or organellar-based phylogenies produced by our analyses (Figs. 4b, c and 5b and Supplementary Figs. 11–15). This observation strongly suggests that host switching on the part of the associated pollinator wasp species, even between distantly related host taxa, has occurred on a regular basis<sup>33</sup> (see Supplementary Notes 3 and 4 for details). Further, each phylogenetic hypothesis makes specific predictions concerning the evolutionary history of sections of *Ficus* and their associations with corresponding genera of pollinator wasps. For example, specific



hypothesized host-switching and hybridization events potentially provide explanations for recent findings of cyto-nuclear incongruence<sup>32–34,38,58</sup>.

Importantly, the frequent pollinator host switches detected throughout *Ficus* evolution potentially provide the mechanism behind the prevalent inferred hybridization. All pollinator host-switching events throughout the coevolution of *Ficus* and agaonid wasps, as inferred through cophylogenetic analyses (between the most strongly supported Astral fig tree and the most recently published wasp tree<sup>20</sup>) have been matched to hybridization events detected between different host fig groups (Supplementary Table 7). Similarly, almost all (18 of 22 listed in Supplementary Table 8) identified hybridization events among main clades of *Ficus* have corresponding host-switching events. These results are also consistent with the inferred pollinator host-switching events and potential genetic exchange in studies of the relationships among closely related sympatric figs in Neotropical section *Americana*<sup>33,34</sup> and in African section *Galoglychia*<sup>32</sup>.

Host-switching as a mechanism underlying hybridization at current ecological timescales is supported by experiments in which pollinator wasps are introduced to figs that are not their natural or most common hosts<sup>37,59,60</sup>. Even among distantly related hosts, hybridization is known to result in viable hybrid seeds, seedlings, and in some cases, adults<sup>21,22,31–34,36,37</sup> (Fig. 2 and Supplementary Data 1). Notably, Ramirez<sup>36</sup> documented hybridization events between introduced *F. religiosa* and native *F. aurea*. These species are included in our genomic and phylogenetic study and belong to different sections within subgenus *Urostigma*. Perhaps more remarkably, he documented viable hybrids between *F. religiosa* and *F. septica*, which belong to different subgenera (the latter belonging to subgenus *Sycomorus*), and exhibit different sexual systems (monoecious and dioecious, respectively). Both cases studied by Ramirez involved visits by pollinators of a native fig species into figs of another species introduced into its range by humans (*F. religiosa*)<sup>36</sup>. These and other<sup>31,37</sup> examples support the plausibility of our phylogenetic and genomic inferences of hybridization between distinct figs from these same sections and subgenera (Fig. 2 and Supplementary Data 1).

Further exploration of the mechanisms of hybridization requires more evidence on how pollinators recognize the signaling of hosts. It is well established that the “appropriate” pollinator wasp species usually recognize and pollinate the “appropriate” host fig species (those in which they are able to reproduce)<sup>61</sup>. Increasingly, it is recognized that receptive figs produce distinctive blends of volatile compounds that mediate the degree of pollinator-specificity that is observed, usually contributing to reproductive isolation of *Ficus* species<sup>25,39,62</sup>. Determining what compounds (or combinations of compounds) of the host volatile chemical blends are recognized by the wasps and play important role in host recognition is critical<sup>39,63,64</sup>. What is known is that these blends differ in compound composition or in relative proportions of a large suite of shared compounds, many of which serve as well-characterized pollinator attractants across flowering plants<sup>64,65</sup>. What is also known is that fig wasps of a given species sometimes arrive, pollinate, and oviposit in the figs of hosts that are neither their most commonly associated host, nor even necessarily a particularly closely related one<sup>22,28–31,36,37</sup>.

Importantly, the interplay of the host chemical signals and the wasps’ recognition of these signals fundamentally directs gene flow among conspecific and sometimes hetero-specific fig individuals by determining what the wasp recognizes as an “appropriate” host<sup>22,30,61</sup>. One of our previous studies<sup>22</sup> included *F. hainanensis* (one of the 15 species sampled in the present study) and four other figs of subgenus *Sycomorus*. We found that those fig species with the highest rates of “mistake visits” to their

receptive figs by wasp species other than their most common pollinator, emitted volatile blends that were similar to those of the “common” hosts of those “mistaken” wasps. We also found that hybrid trees occurred most frequently between those fig species with the most similar chemical blends. Given that the genetics underlying both fig volatile production and wasp volatile recognition are directly affected by the wasps’ choice of host, the potential for hybrid figs to exhibit markedly different volatile chemical signatures is not only possible, but likely<sup>66</sup>. In addition, shifts in volatile chemistry produced by hybrids and resulting interactions with the wasps that they subsequently attract potentially initiate rapidly evolving feedbacks between the hosts and their pollinators. These feedbacks could isolate hybrids from parental lineages, and contribute to the generation of additional species diversity in both host and wasp<sup>14,15,66–68</sup>. How hybridization influences blends of volatile compounds, and how pollinator fidelity to a new host is established, as well as how chemical signaling by the host figs and recognition by the associated wasp pollinators vary over geographical scales, are critical gaps in our understanding of how hybridization might drive diversification of fig–fig wasp systems. As in a case reported in orchids<sup>68</sup>, hybrid figs likely create new volatile blends, with the subsequent establishment of new pollinator specificity via additional pollinator host-switches.

Our results also offered insights into the temporal and biogeographical reconstructions of the fig–wasp mutualism. All current evidence relevant to *Ficus* biogeographic history indicates that figs and wasps have been associated in a pollination mutualism for a long time (roughly 60–90 Mya)<sup>20,23,69</sup>. Some evidence suggests that the mutualism originated in Eurasia<sup>20</sup>, whereas other evidence suggests a Gondwanan origin<sup>23,69</sup>. All current genetic evidence and all analyses point to one or the other of the two monoecious clades of figs (subgenera *Urostigma* and *Pharmacosycea*) as being sister to the rest (mainly gynodioecious groups) and thus to a single derived origin of gynodioecy in figs<sup>39</sup>. The Astral species tree of figs based on nuclear genomes, and one of the wasp phylogenetic reconstructions<sup>20</sup>, are most consistent with a scenario of an Old World origin, with the extant subgenus *Urostigma* as the basal clade followed by divergence of subgenus *Pharmacosycea* (Old World section *Oreosycea* and New World section *Pharmacosycea*), and the subsequent dispersal of the latter section to what is now the New World. This scenario was also supported by the topology of *Ficus* inferred with PhyloNetworks (Supplementary Fig. 6). According to this scenario, the basal divergence of subgenera *Urostigma* and *Pharmacosycea* was followed by the divergence of the Old World gynodioecious species, and a colonization of the monoecious fig species by wasps that were associated with gynodioecious species (Figs. 1 and 5b).

Alternatively, the PCT tree, both organellar phylogenies, and one of the earlier wasp phylogenies<sup>23</sup>, are more consistent with section *Pharmacosycea* (currently New World) and its associated pollinators (the passively pollinating genus *Tetrapus*) being sister to the rest of the figs and the rest of the pollinator wasps, respectively. Under this scenario, a Gondwanan origin would have been followed by a separation of this currently New World group from the rest, with a later colonization of the New World by a group of strangler figs (section *Americana*) that are most closely related to the African section *Galoglychia*. This scenario also differs from the other in being compatible with the position of the passively pollinating *Tetrapus* wasps as being basal to all other pollinating wasps<sup>23</sup>. Importantly, both scenarios clearly indicate a single origin of gynodioecious species from the monoecious fig clades (subgenus *Urostigma* or *Pharmacosycea*)<sup>39</sup>.

Additional genomic data from subgenera *Urostigma* (particularly from the section *Malvanthera* from Australia) and *Pharmacosycea* and their associated pollinators, as well as fossil

evidence (e.g., from South America, Antarctica, Australia, and India), will be needed to more clearly resolve the issue of which of these current taxa represent the basal groups of figs and wasps. Certainly, there is abundant evidence to support hybridization among even distantly related fig taxa that currently co-occur, and at least some evidence for hybridization among groups that currently do not. The latter is most likely to have occurred during geological periods when the different groups did co-occur. Considering the prevalent hybridization within genus *Ficus* inferred here and the challenges that widespread hybridization presents to current standard methods of species tree inference, future work on phylogenetic relationships and biogeographic history of *Ficus* should explicitly consider the potential effects of hybridization.

In conclusion, we suggest that future studies attempting to reconstruct the phylogenetic relationships between figs and their pollinators should not be constrained by the idea that species or even lineages of pollinators have strictly co-radiated with the currently corresponding species or sections of figs. The rampant hybridization that our complete nuclear, mitochondrial, and chloroplast genome data reveal and substantiate helps to explain why resolving fig phylogenies and fig–wasp co-radiation has been so difficult and controversial<sup>20,21,23,38,57,69,70</sup>. In conjunction with the need for improving wasp phylogenies at large taxonomic scales and for wider sampling from key portions of the fig lineages (e.g., the basal separation of *Urostigma* and *Pharmacosycea*), the most relevant questions shift from “Which lineages of wasps are associated with which lineages of figs?” to “What parts of the fig genome have descended together and what parts have been introduced during different hybridization events?” The timing, location, and genetic consequences of hybridization events, and their roles in adaptation, trait evolution and diversification of the two partners, should become the targets of more in-depth studies. In short, little in the evolution of figs and their wasps appears to make sense except in the light of hybridization and introgression (mediated by pollinator host-switches).

Similar diversification scenarios, with both pollinator specificity-mediated reproductive isolation and pollinator host switch-mediated hybridization, have potentially occurred in other “strict-sense” obligate pollination systems. For example, in Phyllanthaceae-*Epicephala* moth systems, sexually deceptive orchids, and oil-collecting bee–orchid systems, plants are also characterized by high, but not strict-sense, pollinator specificity and by relatively weak post-mating isolation mechanisms, as well as by frequent pollinator host-switching and hybridization<sup>14,16,18</sup>. We suggest that the genomic patterns of hybridization and introgression that we have documented in *Ficus* will help guide future studies of the processes that influence diversification in other obligate and generalist plant-pollinator systems.

## Methods

**Study design.** The analytical frameworks applied included: (1) whole-genome short-read sequencing of 17 species representing all major *Ficus* clades and two outgroups, mapping the reads to the reference genome assembly (*F. microcarpa*)<sup>39</sup>, and de novo assembly of chloroplast and mitochondrial genomes based on the same set of samples; (2) inference of phylogenies of the three genomic compartments of *Ficus*, and reconstruction of biogeographic history; (3) detection of hybridization with ABBA-BABA *D*-statistics, PhyloNetworks and, BUCKy based on nuclear genomic data; (4) reconciliation of phylogenetic incongruence among nuclear, chloroplast, and mitochondrial genomes, and inference of the potential evolutionary events causing the incongruence by mapping of the geographic distribution of species and haplotypes based on a previously published 59-species chloroplast genomic phylogeny<sup>38</sup>; (5) reconciliation of phylogenetic incongruence between current fig phylogenies and the two most widely accepted published pollinator phylogenies<sup>20,23</sup>, to infer the relative contributions of pollinator host-switching events and cospeciation events to fig diversification.

Fifteen fig species representing all major *Ficus* lineages (14 sections of all six subgenera)<sup>21</sup> and two outgroup species, *Morus alba* and *Antiaris toxicaria*, were sampled in this study (Supplementary Table 1). Samples were obtained from living adult trees from the outdoor plant collection (*Ficus* garden) of Xishuangbanna

Tropical Botanical Garden, Chinese Academy of Sciences (XTBG, CAS), or from their natural habitats nearby (more than 50 *Ficus* species occur naturally in the area<sup>71</sup>) and in other countries (*F. maxima* and *F. adhatodifolia* from Brazil and *F. microdictya* from Papua New Guinea). One species, *F. carica*, was obtained from the Kunming Institute of Botany (KIB, CAS). All trees sampled from the XTBG plant collection originated from seeds or seedlings collected from their natural forest habitat nearby or introduced from other countries. All trees were carefully identified based on morphological traits before sampling. Meanwhile, our sample of species is also taxonomically and geographically well representative based on current classification and distribution (in species and section level)<sup>20,21</sup>. Pollinator phylogenies (Supplementary Fig. 5) were extracted from two published fig wasp phylogenies<sup>20,23</sup>.

**Genome sequencing and SNP genotyping.** Whole genomic DNA was extracted from ~100 mg of silica gel-dried fresh leaf tissue using a modified cetyltrimethylammonium bromide method<sup>72</sup>. A minimum of 5 µg of total genomic DNA from each sample was used to construct an independent sequencing library following the manufacturer’s instructions, before sequencing using the Illumina HiSeq X Ten (Illumina Inc.). To avoid potential DNA contamination during sequencing, such as that via index-swapping, we constructed dual-indexed libraries with unique indexes for each sample. The dual indexes contained a total of 16 base pairs (bp) and were inserted into the flanking regions of the target DNA fragments. Raw reads that had any mismatch with index sequences were clustered as undetermined sequences and removed from our analysis. Meanwhile, the unique dual-index of each sample also ensured unambiguous separation of raw sequenced data and avoidance of potential sequences contamination among samples.

Adaptors and low-quality bases were removed from the raw pair-end reads using Trimmomatic-v0.39 (ref. <sup>73</sup>) after performing quality control with FastQC-v0.11.9 (ref. <sup>74</sup>). The reads were filtered with a sliding window of size 7, with an average Phred score = 20 within each window. The trimmed reads were mapped to a high-quality *F. microcarpa* reference genome<sup>39</sup> assembly with 93.5% BUSCO completeness (submitted to GSA database (<http://gsa.big.ac.cn/>) under accession number GWHABKV00000000, Supplementary Note 1) using BWA-v0.7.17 (ref. <sup>75</sup>) with default parameters. The mapped reads were sorted, and duplicated reads were removed using SAMtools-v1.9 (ref. <sup>76</sup>). The Realigner Target Creator and Indel Realigner programs from the Genome Analysis Toolkit (GATK-v4) package<sup>77</sup> were used for global realignment of reads around indels from the sorted BAM files.

To identify a high-quality SNP dataset, we applied the GATK pipeline following the suggested best practices workflow (<https://gatk.broadinstitute.org/hc/en-us/articles/360036194592-Getting-started-with-GATK4>). The GATK best practices included a number of steps to ensure variant accuracy. To remove erroneous mismatches around small indels, IndelRealigner was applied to process the alignment of BAM files. HaplotypeCaller was used to call variants for each sample and the program GenotypeGVCFs was used to merge all of the individual GVCF (Genomic Variant Call Format) files. The merged GVCF file was subject to quality control so that the minimum quality of each SNP was 30. To avoid false positives of read mapping in repetitive sequences, sequencing depth of each high-quality SNP ranged from 2 to 200. A total of 4,473,644 SNPs were used to construct three non-overlapping size-fixed windows, with window sizes of 50, 100, and 500 kb. A second, more conservative SNPs calling strategy was also used. In this strategy, samtools/bcftools was applied to call SNPs using the same data with default parameters. SNPs that were present in both the GATK and the samtools/bcftools datasets were used for further analysis. In addition, we allowed each SNP site with a maximum missing rate of 40% across all samples tested. This ensured that the two outgroup species (*M. alba* and *A. toxicaria*) supplied abundant information for inference of the phylogenetic tree. A total of 389,834 SNPs in non-repeat regions were extracted from the VCF data, using a home-made PERL script (<https://github.com/tangerzhang/VCFplayer/blob/master/vcf2fasta.pl>). Based on the second SNPs dataset (using samtools/bcftools), we used a sliding-window approach and constructed a SNPs-fixed window with window size of 1000 SNPs and a step size of 1100 SNPs, meaning a space of 100 SNPs between two successive windows. To maintain abundant information within each window in the above four windows datasets, we removed windows with fewer than 50 SNPs, and generated 351, 764, 3697, and 3459 windows for the 1000SNP, 500-kb, 100-kb, and 50-kb datasets, respectively (Supplementary Table 3). The physical location of each window in chromosomes of the reference genome was labeled (Source Data 1). All four datasets were used to infer species trees with ASTRAL-II (ref. <sup>40</sup>); the 1000SNP dataset was also used for BUCKy-v1.4.4 analysis<sup>43</sup>; and the 500-kb datasets were further used for PhyloNetworks analysis<sup>45</sup>.

Both chloroplast and mitochondrial genomes were de novo assembled with the recommended pipeline of the GetOrganelle toolkit<sup>78</sup>. Chloroplast genomes of all 17 species and mitochondrial genomes of 14 species were assembled in circular sequences. For three species whose mitochondrial-genome assemblies were not circular, their assembly graphs were checked using Bandage program<sup>79</sup>, and non-target contigs were manually cleaned. The cleaned assembly graphs of circular genomes or contigs were exported to FASTA sequences using the GetOrganelle toolkit (Supplementary Table 2 and Appendix 5). SNPs in chloroplast and mitochondrial genomes from the same sequencing raw data were identified using the GATK pipeline as mentioned above. In total, we identified 5486 SNPs in

chloroplast genomes and 4649 SNPs in mitochondrial genomes (Supplementary Table 2). All sampled species are inferred as diploid species based on ploidy estimation with nQuire program<sup>80</sup> (Supplementary Table 9).

**Reconstruction of *Ficus* evolutionary history.** We inferred *Ficus* phylogenies based on the nuclear, chloroplast, and mitochondrial genomes of the target figs with multiple methods separately. We first inferred the Astral species trees under a multi-species coalescent model with the four genomic-windows datasets, using the ASTRAL-II program<sup>40</sup>. For each of the four windows datasets, gene trees of all windows were inferred using the RAxML-v8 program<sup>81</sup> with GTR+GAMMA substitution model and 1000 bootstrap replications. The Astral species tree was inferred with greedy consensus<sup>40</sup> of the 1000 bootstrapped replicate trees of each gene tree, and both support values (showing the percentage of bootstrap replicates that contain a given branch) and branch length in coalescent units were inferred (Supplementary Fig. 1).

We also inferred the PCT tree of *Ficus* under the BCA-based<sup>42</sup> BUCKy<sup>43</sup> with the 1000SNP dataset. Whereas the coalescent-based Astral tree assumes that ILS is the sole reason for incongruence in gene trees, the BCA method makes no such assumption and allows both ILS and hybridization as causes of incongruence. The PCT tree with the highest CF value was reconstructed to summarize the main vertical phylogenetic signal (i.e., the best-supported tree branching pattern) of all gene trees based on all single windows. The CF value gives the proportion of genomic windows (or genes) supporting a given clade<sup>44</sup>. In brief, we first ran Bayesian analysis of each individual gene (window) with MrBayes-v.3.2.5 (ref. <sup>82</sup>) to obtain a posterior distribution of gene trees. For each gene tree inference, under the GTR+GAMMA substitution model, two Markov chain Monte Carlo simulations were run with one cold chain and three heated chains for chain lengths of 1,000,000 (10% burn-ins) and sampling every 100 generations. Finally, we used BUCKy to estimate the PCT tree and CF values of different splits.

*Ficus* divergence time was estimated using the Bayesian inference in the MCMCTREE program of PAML 4.9a<sup>83</sup>, based on the topology of the Astral *Ficus* species tree using the 100-kb dataset, which was for the most part corroborated by the other Astral trees (Supplementary Fig. 1). MCMCTREE was used to estimate the branch length with datasets of 10 partitions of the whole aligned genome<sup>84</sup> (Appendix 3). To decide the rate prior, we first ran BASEML to estimate the overall substitution rate under a strict molecular clock with the root age (Crown Moraceae) set to 93 Mya<sup>85</sup> for a rough estimation. We set  $r_{\text{gene}} \gamma = 1, 2$ , with average  $0.5 \times 10^{-8}$  substitutions per site per year. We then estimated the divergence times and 95% credible interval for each node with approximate likelihood principle<sup>86</sup> under HKY85 model, and with independent rates model (clock = 2). Seven fossil calibrations following contemporary accepted standards on Moraceae and *Ficus*<sup>20,85</sup> (Supplementary Table 4) were placed on the phylogeny using soft bounds<sup>87</sup>. The MCMC analyses ran 200,000,000 iterations with sampling every 10,000 cycles after a burn-in of 200,000. The analyses were run twice to ensure chain convergence with the effective sample size (ESS) larger than 350, using Tracer v.1.6 (<http://beast.bio.ed.ac.uk/Tracer>).

To investigate the general biogeographical history of *Ficus*, we further reconstructed the ancestral geographic distribution of *Ficus* lineages using the R package “BioGeoBEARS”<sup>88</sup>. We categorized current species distributions into four states, Eurasia, Afrotropics, Australasia, and Neotropics, following previous studies<sup>20</sup>. We inferred the ancestral distribution probabilities of *Ficus* lineages based on the present distribution states of the 15 fig species, and the Astral species tree with fossil calibration (Supplementary Fig. 2). Maximum number of regions (of the four listed above) in which a lineage could be distributed was set as two. Six biogeographical models were used, including three models allowing jump dispersal events<sup>88</sup>. Choice of the final pattern of *Ficus* biogeographical history retained for analysis was based on results of the best-fit model (Supplementary Table 10 and Appendix 4). Following the same methodology, divergence time and ancient biogeography of *Ficus* plants were also inferred based on the PCT tree.

Phylogenies of the two organellar genomes of *Ficus* were also inferred separately based on the concatenated full genomic sequences (Appendix 5). The full sequences were automatically aligned in MAFFT version 7.0 (ref. <sup>89</sup>) using the default option. The ML tree of chloroplast genomes was reconstructed using RAxML<sup>79</sup> with the most appropriate substitution model (GTR+GAMMA) inferred using jmodelTest2 (ref. <sup>90</sup>) and 1000 bootstrap replications, which is the same strategy that was used in a similar previous study<sup>39</sup>. Mitochondrial genomes showed high rates of reversion and translocations. The sequence orders were adjusted in MULAN (<https://mulan.dcode.org/>) using *Ficus adhatodifolia* (WG105) as the reference, aligned in MAFFT, then manually adjusted using Geneious version 11 (<https://www.geneious.com>). The alignment was trimmed using 10% gapping option for sites. The ML tree of the mitochondrial genomes was reconstructed using FastTree-2 with the GTR substitution model and a Gamma20-based likelihood option under an approximately-maximum-likelihood method<sup>91</sup>.

**Distinguishing hybridization from ILS.** First, we used ABBA-BABA *D*-statistics<sup>46,47</sup> to detect hybridization along a specific branch of the phylogeny, based on the frequencies of discordant SNP genealogies in a four-taxon tree. The method is based on the relative abundance of two allelic configurations called “ABBA” or “BABA”. Given four taxa, P1, P2, P3, and an outgroup O, with the relationship “[(P1, P2), P3], O”, ABBA patterns are SNPs where P1 retains the outgroup allele

and P2 and P3 share the derived allele. Similarly, BABA patterns are SNPs where P1 and P3 share the derived allele and P2 retains the outgroup allele. Under the null hypothesis of no gene flow, the ABBA and BABA patterns should be equally frequent via ILS and the value of *D* should be equal to zero. A significantly positive *D* value indicates gene flow between P2 and P3, whereas a significantly negative *D* value indicates gene flow between P1 and P3 (ref. <sup>47</sup>). Here, an improved *D*-statistic<sup>46</sup> method calculated with the program doAbbababa2 implemented in ANGSD<sup>92</sup> was used. Briefly, BAM files that were used for variant detection of all 17 species sampled were recruited for this analysis. All samples underwent error correction and ABBA-BABA *D*-statistics analysis was conducted at different taxonomic/phylogenetic levels according to the instructions on the ANGSD website (<http://www.popgen.dk/angsd/index.php/Abbababa2>). With these capacities, the improved *D*-statistics can obtain robust results even when sequencing depth is very low (1–10X)<sup>46</sup>.

To detect hybridization events throughout *Ficus* evolution, we conducted ABBA-BABA *D*-statistics for four-taxon groups hierarchically, from the species tips to the main clades following the topology of the Astral species tree. When *D*-statistics analysis was conducted in four-taxon groups at higher taxonomic/phylogenetic levels, all species within each taxon (section, subgenus, or main clade) were treated as sample replications (Supplementary Data 1). Only monophyletic higher-level clades or subgenera in Astral species trees were used (Supplementary Fig. 1). A non-*Ficus* species (*Morus alba*) was taken as outgroup in all tested four-taxon groups finally shown. The doAbbababa2 program could calculate the *D*-statistics for all potential four-taxon combinations of species or groups. This means that for each species or clade, one could potentially infer multiple hybridization events with other taxa if the species was included in multiple four-taxon groups. However, only those four-taxon groups that followed the phylogenetic topology of the Astral species tree were considered. In a few cases, results of some alternative four-taxon groups were also kept when the phylogenetic topology among taxa was unstable (Supplementary Data 1).

In addition, we inferred *Ficus* hybridization events using PhyloNetworks<sup>45</sup>, which infers phylogenetic networks with maximum pseudolikelihood from gene trees or quartet CF value. It allows both vertical inheritance under the coalescent model, and horizontal inheritance (e.g. hybridization) with reticulation nodes in the network. Following the online pipeline (<https://github.com/crsl4/PhyloNetworks.jl/wiki>), PhyloNetworks analyses were conducted with the 500-kb windows dataset. *Ficus* networks were inferred based on the ML gene trees generated with the 500-kb windows dataset. We ran SNaQ to search for the best network with the maximum hybrid node number (hmax) increasing from zero to 15, each hmax being repeated ten times. The optimal number of hybridization events was chosen when the negative log pseudolikelihood (–logplik) score of an additional hmax did not significantly decrease.

Furthermore, we also inferred *Ficus* hybridization events based on the analysis of results under BUCKy on inference of species trees based on the 1000SNP dataset (Appendix 2). The BCA-based species tree inference using BUCKy makes no particular assumption about possible causes of gene tree incongruence (e.g., ILS, hybridization, horizontal gene transfer). With the similar function of above ABBA-BABA *D*-statistics and PhyloNetworks, BUCKy provides a third way to distinguish hybridization from ILS by testing whether the distribution of gene-tree frequencies is compatible with that expected under neutral coalescent processes (ILS), or whether other biological processes must also be invoked (e.g., hybridization, horizontal gene transfer)<sup>44,48</sup>. For example, under a coalescent process, for a three-taxon group ((a,b),c), the most probable topology is (a,b),c, i.e., the topology that matches the species tree and has the highest frequency; and the two non-matching alternative minor topologies, (a,c),b and (b,c),a, are expected to be equally frequent (ref. <sup>48</sup>). However, if the frequencies of the two non-matching minor topologies are significantly different, the coalescent model can be rejected and the contribution to gene-tree incongruence of other biological processes additional to ILS, such as hybridization or horizontal gene transfer, can be inferred<sup>44,48</sup>. Under BUCKy, the CF value and its credibility interval for each clade can be estimated, indicating the proportion of genes supporting the clade<sup>44</sup>. The clades with the largest CF values produce the PCT tree, and non-overlapping CF credibility intervals of two corresponding alternative minor clades can be used as evidence of hybridization in addition to ILS.

**Cophylogenetic analyses and detection of associate switches.** Phylogenetic incongruences between nuclear and organellar genomes of *Ficus*, and between figs and their pollinating wasps, could result from hybridization and pollinator host switches. Event-based cophylogenetic reconciliation analyses were conducted using the program JANE 4.0 (ref. <sup>49</sup>) to infer phylogenetic incongruences based on five types of evolutionary events: co-divergence/cospeciation, duplication, associate switches, associate losses, and failure to diverge. Incongruence between nuclear and organellar genomes, especially the associate (host-)switch events, indicates organellar genome exchange among *Ficus* taxa with different nuclear genomes, indirectly supporting the occurrence of hybridization between *Ficus* species. Host-switch events inferred between fig and pollinator phylogenies indicate pollinator host switches among *Ficus* taxa. Considering the close biological relationship between the fig/pollinator wasp interaction and *Ficus* hybridization, detection of corresponding host switches in both organellar genome and pollinator wasps between the same pair of *Ficus* taxa would strongly suggest that pollinator host-switches generate *Ficus* hybridization. The cophylogenetic reconciliations included

those between all pairs of fig phylogenies based on nuclear, chloroplast, and mitochondrial genomes of the 15 fig species generated here, and phylogenies of their associated pollinating wasps extracted from previous studies<sup>20,23</sup>. JANE reconciles two phylogenies by minimizing the total costs with a certain cost setting for each of the five kinds of events listed above. Results of the cophylogenetic analyses depend on the cost weights assigned to different evolutionary events. We thus first selected the optimal cost setting by running both random tip mapping and random parasite tree permutation tests, with the cost range of each evolutionary event varying from 0 to 1 (except for associate switch events, which vary from 0 to 2). The genetic algorithm parameters were set following a previous study<sup>20</sup>, with generation number = 40, population size = 1000 and sample size = 100. The optimal cost settings, with *P*-value differing significantly from the random setting and with lowest observed cost value, were chosen for formal cophylogenetic analyses (Supplementary Table 8), which were then run in solve mode to find the best solutions to reconcile each pair of phylogenies under optimal cost settings and using the same genetic algorithm parameters.

Results of cophylogenetic analyses are highly dependent on the phylogenetic topology. To obtain more robust results, two alternative nuclear trees (Astral species tree and PCT tree) (Supplementary Fig. 1B, E), ML trees of the chloroplast and mitochondrial genomes (Supplementary Fig. 4), and two alternative ML trees of associated pollinator wasps (Supplementary Fig. 5) were used. Although neither of the phylogenies of pollinator wasps is as robust as those for *Ficus* generated here, they represent two alternative views of fig wasp evolution that are both considered plausible<sup>20,23</sup>. The most notable difference between the two pollinator phylogenies is the identity of the earliest-diverging wasp clade. The pollinator phylogeny of Machado et al.<sup>23</sup> treated *Tetrapus* wasps (whose hosts are the Neotropical section *Pharmacosycea*) as sister to all other pollinating wasps, a placement that better matches the PCT tree of *Ficus*, which places section *Pharmacosycea* as sister to all other figs. This latter pair of phylogenies (Machado et al.<sup>23</sup> pollinators tree, *Ficus* PCT tree) reflects a long-held view of the evolutionary history of fig-fig wasp mutualism, notably the hypothesis that passive pollination is the primitive condition in figs and wasps. In contrast, in the pollinator phylogeny published by Cruaud et al.<sup>20</sup>, *Tetrapus* was inferred to lie in the middle of the wasp phylogeny, while section *Pharmacosycea* was also inferred to lie near the middle of the Astral *Ficus* species tree (Supplementary Note 3).

#### Geographic clustering pattern in the chloroplast phylogeny of 59 fig species.

If hybridization is the main reason for nuclear-organellar genome incongruence, we could expect that fig species with long histories of geographic co-occurrence may have undergone exchange of organellar genomes during hybridization. This would lead to geographic clustering of similar chloroplast genomes. To test this hypothesis, we surveyed 59 fig species for which a chloroplast genome phylogeny<sup>38</sup> and information on geographic distribution were both available (Supplementary Data 3 and Appendix 9). These represented all six subgenera and most main sections. Taxonomic data and the distribution area of each species represented in both nuclear and chloroplast phylogenies were collected<sup>20,21,38</sup> and mapped onto the nuclear and chloroplast trees using the R package *ape*-v5.3<sup>93</sup>, to investigate the hypothesized geographic clustering patterns in the chloroplast phylogeny.

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

#### Data availability

All data needed to evaluate the conclusions in the paper are presented in the paper, and its Supplementary information files. Reporting summary for this article is available as a Supplementary information file. All raw resequencing data produced for this project have been deposited at the Genbank with the BioProject number as PRJNA684963, and the Genome Sequence Archive BIG Data Center (<http://bigd.big.ac.cn/gsa>) under the accession number of CRA001150 separately. All other raw results related to all figures or tables are supplied as Appendix files, which are available at Dryad (<https://doi.org/10.5061/dryad.zcrjdfn7m>)<sup>94</sup>. Source data are provided with this paper.

#### Code availability

All codes needed to evaluate the conclusions in the paper are presented in the paper, Source Data, and GitHub (<https://doi.org/10.5281/zenodo.4308886>)<sup>95</sup>.

Received: 26 September 2019; Accepted: 4 January 2021;

Published online: 02 February 2021

#### References

- Taylor, S. A. & Larson, E. L. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nat. Ecol. Evol.* **3**, 170–177 (2019).

- Payseur, B. A. & Rieseberg, L. H. A genomic perspective on hybridization and speciation. *Mol. Ecol.* **25**, 2337–2360 (2016).
- Arnold, M. L. & Kunte, K. Adaptive genetic exchange: a tangled history of admixture and evolutionary innovation. *Trends Ecol. Evol.* **32**, 601–611 (2017).
- Mallet, J. Hybrid speciation. *Nature* **446**, 279–283 (2007).
- Abbott, R. et al. Hybridization and speciation. *J. Evol. Biol.* **26**, 229–246 (2013).
- Gross, B. L. & Rieseberg, L. H. The ecological genetics of homoploid hybrid speciation. *J. Hered.* **96**, 241–252 (2005).
- Schumer, M., Rosenthal, G. G. & Andolfatto, P. How common is homoploid hybrid speciation? *Evolution* **68**, 1553–1560 (2014).
- Grant, V. Pollination systems as isolating mechanisms in angiosperms. *Evolution* **3**, 82–97 (1949).
- Kay, K. M. & Sargent, R. D. The role of animal pollination in plant speciation: Integrating ecology, geography, and genetics. *Annu. Rev. Ecol. Syst.* **40**, 637–656 (2009).
- Serrano-Serrano, M. L., Rolland, J., Clark, J. L., Salamin, N. & Perret, M. Hummingbird pollination and the diversification of angiosperms: an old and successful association in Gesneriaceae. *Proc. R. Soc. B Biol. Sci.* **284**, <https://doi.org/10.1098/rspb.2016.2816> (2017).
- Thompson, J. N. Specific hypotheses on the geographic mosaic of coevolution. *Am. Nat.* **153**, S1–S14 (1999).
- Van der Niet, T., Peakall, R. & Johnson, S. D. Pollinator-driven ecological speciation in plants: new evidence and future perspectives. *Ann. Bot.* **113**, 199–211 (2014).
- Armbruster, W. S. The specialization continuum in pollination systems: diversity of concepts and implications for ecology, evolution and conservation. *Funct. Ecol.* **31**, 88–100 (2017).
- Ayasse, M., Stokl, J. & Francke, W. Chemical ecology and pollinator-driven speciation in sexually deceptive orchids. *Phytochemistry* **72**, 1667–1677 (2011).
- Machado, C. A., Robbins, N., Gilbert, M. T. P. & Herre, E. A. Critical review of host specificity and its coevolutionary implications in the fig/fig-wasp mutualism. *Proc. Natl Acad. Sci. USA* **102**, 6558–6565 (2005).
- Kawakita, A. Evolution of obligate pollination mutualism in the tribe Phyllanthaeae (Phyllanthaceae). *Plant Species Biol.* **25**, 3–19 (2010).
- Ramirez, W. Host specificity of fig wasps (Agaonidae). *Evolution* **24**, 680–691 (1970).
- Schiestl, F. P. & Schluter, P. M. Floral isolation, specialized pollination, and pollinator behavior in orchids. *Annu. Rev. Entomol.* **54**, 425–446 (2009).
- Ramirez, S. R. et al. Asynchronous diversification in a specialized plant-pollinator mutualism. *Science* **333**, 1742–1746 (2011).
- Cruaud, A. et al. An extreme case of plant-insect co-diversification: figs and fig-pollinating wasps. *Syst. Biol.* **61**, 1029–1047 (2012).
- Berg, C. C. & Corner, E. J. H. In *Flora Malesiana Series I - Seed Plants* Vol. 17 (ed. Nootboom, H. P.) 1–702 (Nationaal Herbarium, Nederland, 2005).
- Wang, G., Cannon, C. H. & Chen, J. Pollinator sharing and gene flow among closely related sympatric dioecious fig taxa. *Proc. R. Soc. B Biol. Sci.* **283**, <https://doi.org/10.1098/rspb.2015.2963> (2016).
- Machado, C. A., Jouselin, E., Kjellberg, F., Compton, S. G. & Herre, E. A. Phylogenetic relationships, historical biogeography and character evolution of fig-pollinating wasps. *Proc. R. Soc. B Biol. Sci.* **268**, 685–694 (2001).
- Harrison, R. D. Figs and the diversity of tropical rainforests. *Bioscience* **55**, 1053–1064 (2005).
- Grison-Pigé, L., Bessière, J. M. & Hossaert-McKey, M. Specific attraction of fig-pollinating wasps: Role of volatile compounds released by tropical figs. *J. Chem. Ecol.* **28**, 283–295 (2002).
- Herre, E. A. et al. Molecular phylogenies of figs and their pollinator wasps. *J. Biogeogr.* **23**, 521–530 (1996).
- Molbo, D., Machado, C. A., Sevenster, J. G., Keller, L. & Herre, E. A. Cryptic species of fig-pollinating wasps: Implications for the evolution of the fig-wasp mutualism, sex allocation, and precision of adaptation. *Proc. Natl Acad. Sci. USA* **100**, 5867–5872 (2003).
- Rasplus, J. Y. in *The Biodiversity of African Plants* (eds van der Maesen, L. J. G. et al.) 639–649 (Springer, 1996).
- Yang, L.-Y. et al. The incidence and pattern of co-pollinator diversification in dioecious and monoecious figs. *Evolution* **69**, 294–304 (2015).
- Cornille, A. et al. Floral volatiles, pollinator sharing and diversification in the fig-wasp mutualism: insights from *Ficus natalensis*, and its two wasp pollinators (South Africa). *Proc. R. Soc. B Biol. Sci.* **279**, 1731–1739 (2012).
- Compton, S. G. A collapse of host specificity in some African fig wasps. *S. Afr. J. Sci.* **86**, 39–40 (1990).
- Renoult, J. P., Kjellberg, F., Grout, C., Santoni, S. & Khadari, B. Cyto-nuclear discordance in the phylogeny of *Ficus* section *Galoglychia* and host shifts in plant-pollinator associations. *BMC Evol. Biol.* **9**, 248 (2009).
- 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).

34. Jackson, A. P., Machado, C. A., Robbins, N. & Herre, E. A. Multi-locus phylogenetic analysis of neotropical figs does not support co-speciation with the pollinators: the importance of systematic scale in fig/wasp cophylogenetic studies. *Symbiosis* **45**, 57–72 (2008).
35. Parrish, T. L., Koelewijn, H. P., van Dijk, P. J. & Kruijt, M. Genetic evidence for natural hybridization between species of dioecious *Ficus* on island populations. *Biotropica* **35**, 333–343 (2003).
36. Ramirez, W. Hybridization of *Ficus religiosa* with *F. septica* and *F. aurea* (Moraceae). *Rev. Biol. Trop.* **42**, 339–342 (1994).
37. Wei, Z. D., Kobmoo, N., Cruaud, A. & Kjellberg, F. Genetic structure and hybridization in the species group of *Ficus auriculata*: can closely related sympatric *Ficus* species retain their genetic identity while sharing pollinators? *Mol. Ecol.* **23**, 3538–3550 (2014).
38. Bruun-Lund, S., Clement, W. L., Kjellberg, F. & Rønsted, N. First plastid phylogenomic study reveals potential cyto-nuclear discordance in the evolutionary history of *Ficus* L. (Moraceae). *Mol. Phylogenet. Evol.* **109**, 93–104 (2017).
39. Zhang, X. et al. Genomes of the Banyan tree and pollinator wasp provide insights into fig-wasp coevolution. *Cell* **183**, 875–889 (2020).
40. Mirarab, S. & Warnow, T. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics* **31**, 44–52 (2015).
41. Rønsted, N., Weiblen, G. D., Clement, W. L., Zerega, N. J. C. & Savolainen, V. Reconstructing the phylogeny of figs (*Ficus*, Moraceae) to reveal the history of the fig pollination mutualism. *Symbiosis* **45**, 45–55 (2008).
42. Ane, C., Larget, B., Baum, D. A., Smith, S. D. & Rokas, A. Bayesian estimation of concordance among gene trees. *Mol. Biol. Evol.* **24**, 412–426 (2007).
43. Larget, B. R., Kotha, S. K., Dewey, C. N. & Ane, C. BUCKy: Gene tree/species tree reconciliation with Bayesian concordance analysis. *Bioinformatics* **26**, 2910–2911 (2010).
44. Baum, D. A. Concordance trees, concordance factors, and the exploration of reticulate genealogy. *Taxon* **56**, 417–426 (2007).
45. Solis-Lemus, C., Bastide, P. & Ane, C. PhyloNetworks: a package for phylogenetic networks. *Mol. Biol. Evol.* **34**, 3292–3298 (2017).
46. Soraggi, S., Wiuf, C. & Albrechtsen, A. Powerful inference with the D-statistic on low-coverage whole-genome data. *G3 (Bethesda)* **8**, 551–566 (2018).
47. Durand, E. Y., Patterson, N., Reich, D. & Slatkin, M. Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* **28**, 2239–2252 (2011).
48. Degnan, J. H. & Rosenberg, N. A. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* **24**, 332–340 (2009).
49. Conow, C., Fielder, D., Ovadia, Y. & Libeskind-Hadas, R. Jane: a new tool for the cophylogeny reconstruction problem. *Algorithms Mol. Biol.* **5**, <https://doi.org/10.1186/1748-7188-5-16> (2010).
50. Ramsey, A. J. & Mandel, J. R. When one genome is not enough: organellar heteroplasmy in plants. *Annual Plant Reviews* **2**, 619–658 (2019).
51. Zhang, Q. & Liu, Y. & Sodmergen. Examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 angiosperm species. *Plant Cell Physiol.* **44**, 941–951 (2003).
52. Hu, Y. C., Zhang, Q. & Rao, G. Y. & Sodmergen. Occurrence of plastids in the sperm cells of Caprifoliaceae: Biparental plastid inheritance in angiosperms is unilaterally derived from maternal inheritance. *Plant Cell Physiol.* **49**, 958–968 (2008).
53. Mayr, E. *Animal Species and Evolution* 1–811 (Belknap Press, 1963).
54. Wu, C. I. The genic view of the process of speciation. *J. Evol. Biol.* **14**, 851–865 (2001).
55. Sun, M. et al. Deep phylogenetic incongruence in the angiosperm clade Rosidae. *Mol. Phylogenet. Evol.* **83**, 156–166 (2015).
56. Folk, R. A., Soltis, P. S., Soltis, D. E. & Guralnick, R. New prospects in the detection and comparative analysis of hybridization in the tree of life. *Am. J. Bot.* **105**, 364–375 (2018).
57. Jiao, X., Flouri, T., Rannala, B. & Yang, Z. The impact of cross-species gene flow on species tree estimation. *Syst. Biol.* **69**, 830–847 (2020).
58. Jouselin, E. et al. One fig to bind them all: host conservatism in a fig wasp community unraveled by cospeciation analyses among pollinating and nonpollinating fig wasps. *Evolution* **62**, 1777–1797 (2008).
59. Moe, A. M. & Weiblen, G. D. Pollinator-mediated reproductive isolation among dioecious fig species (*Ficus*, Moraceae). *Evolution* **66**, 3710–3721 (2012).
60. Wang, G., Compton, S. G. & Chen, J. The mechanism of pollinator specificity between two sympatric fig varieties: a combination of olfactory signals and contact cues. *Ann. Bot.* **111**, 173–181 (2013).
61. Bronstein, J. L. Maintenance of species-specificity in a neotropical fig-pollinator wasp mutualism. *Oikos* **48**, 39–46 (1987).
62. Ware, A., Kaye, P., Compton, S. & Noort, S. Fig volatiles: their role in attracting pollinators and maintaining pollinator specificity. *Plant Syst. Evol.* **186**, 147–156 (1993).
63. Soler, C. C. L., Proffitt, M., Bessière, J. M., Hossaert-McKey, M. & Schatz, B. Evidence for intersexual chemical mimicry in a dioecious plant. *Ecol. Lett.* **15**, 978–985 (2012).
64. Hossaert-McKey, M., Soler, C., Schatz, B. & Proffitt, M. Floral scents: their roles in nursery pollination mutualisms. *Chemoecology* **20**, 75–88 (2010).
65. Knudsen, J. T., Eriksson, R., Gershenzon, J. & Stahl, B. Diversity and distribution of floral scent. *Bot. Rev.* **72**, 1–120 (2006).
66. Herre, E. A., Jander, 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).
67. Kiester, A. R., Lande, R. & Schemske, D. W. Models of coevolution and speciation in plants and their pollinators. *Am. Nat.* **124**, 220–243 (1984).
68. Vereecken, N. J., Cozzolino, S. & Schiestl, F. P. Hybrid floral scent novelty drives pollinator shift in sexually deceptive orchids. *BMC Evol. Biol.* **10**, 103 (2010).
69. Rønsted, N. et al. 60 million years of co-divergence in the fig-wasp symbiosis. *Proc. R. Soc. B Biol. Sci.* **272**, 0962–8452 (2005). 2593–2599.
70. Wiebes, J. T. Co-evolution of figs and their insect pollinators. *Annu. Rev. Ecol. Syst.* **10**, 1–12 (1979).
71. Zhu, H. et al. *Native Seed Plants in Xishuangbanna of Yunnan* (eds Zhu, H. & Yan, L.) 1–565 (Science Press, 2012).
72. Yang, J. B., Li, D. Z. & Li, H. T. Highly effective sequencing whole chloroplast genomes of angiosperms by nine novel universal primer pairs. *Mol. Ecol. Resour.* **14**, 1024–1031 (2014).
73. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
74. Andrews, S. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (2010).
75. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
76. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
77. McKenna, A. et al. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
78. Jin, J.-J. et al. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* **21**, 31 (2020).
79. Wick, R. R., Schultz, M. B., Zobel, J. & Holt, K. E. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* **31**, 3350–3352 (2015).
80. Weiß, C. L., Pais, M., Cano, L. M., Kamoun, S. & Burbano, H. A. nQuire: a statistical framework for ploidy estimation using next generation sequencing. *BMC Bioinformatics* **19**, 122 (2018).
81. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
82. Ronquist, F. et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
83. Yang, Z. H. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
84. Zhu, T. Q., Dos Reis, M. & Yang, Z. H. Characterization of the uncertainty of divergence time estimation under relaxed molecular clock models using multiple loci. *Syst. Biol.* **64**, 267–280 (2015).
85. Gardner, E. M., Sarraf, P., Williams, E. W. & Zerega, N. J. C. Phylogeny and biogeography of *Maclura* (Moraceae) and the origin of an anachronistic fruit. *Mol. Phylogenet. Evol.* **117**, 49–59 (2017).
86. dos Reis, M. & Yang, Z. Approximate likelihood calculation on a phylogeny for bayesian estimation of divergence times. *Mol. Biol. Evol.* **28**, 2161–2172 (2011).
87. Yang, Z. & Rannala, B. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* **23**, 212–226 (2006).
88. Matzke, N. J. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in Island Clades. *Syst. Biol.* **63**, 951–970 (2014).
89. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
90. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9**, 772–772 (2012).
91. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
92. Kornelissen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 1–13 (2014).
93. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
94. Wang, G. et al. Data from: Genomic evidence of prevalent hybridization throughout the evolutionary history of the fig-wasp pollination mutualism. *Dryad, Dataset* <https://doi.org/10.5061/dryad.zcrjdfn7m> (2020).

95. Zhang, T. & Zhang, S. C. Code from: Genomic evidence of prevalent hybridization throughout the evolutionary history of the fig-wasp pollination mutualism. *GitHub* <https://doi.org/10.5281/zenodo.4308886> (2020).

### Acknowledgements

We thank Bruno DeMedeiros and Egbert Leigh for useful comments, the Xishuangbanna National Nature Reserve, the Xishuangbanna Station for Tropical Rainforest Ecosystem Studies, and the Gardening and Horticulture Department of XTBG, CAS, for allowing sample collection; Simon Segar and Daniel Souto-Vilarós for collecting the *F. microdictya* sample; the Public Technology Service Center of XTBG, CAS, for the use of its DNA extraction platform; and Richard Corlett and XTBG for organizing the “Scientific Paper Writing: XTBG Advanced Workshop”, during which the manuscript was significantly improved. This work was supported by the National Natural Science Foundation of China (grant numbers: 31870359, U1402264, and 31400329), the Yunnan Natural Science Foundation (grant number: 2014FB185), Youth Innovation Promotion Association of CAS (grant number: 2020392), and the “International Partner Program” of CAS (to E.A.H., C.A.M., and M.L.A.). E.A.H. also acknowledges the Smithsonian Tropical Research Institute and Smithsonian’s Scholarly Studies program for financial support.

### Author contributions

G.W. and J.C. designed the study; G.W. and R.A.S.P. collected the samples; G.W., X.Z., W-B.Y., Y.W., and D. Ma conducted the genome data analyses; G.W., X.Z., and J.C. wrote the first draft of the manuscript; G.W., E.A.H., D. McKey, C.A.M., X.Z., W-B.Y., and J.C. wrote the revised manuscript; C.H.C., M.L.A., R.A.S.P., R.M., and Y-F.L. also made significant contributions to the manuscript and contributed to data analysis. All authors gave final approval for publication.

### 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/s41467-021-20957-3>.

**Correspondence** and requests for materials should be addressed to G.W. or J.C.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021