Bursts of transposable elements and adaptive evolution of fig wasps

Jing Liu[1], Yun-Heng Miao[1], Da-Wei Huang[1], and Jinhua Xiao[1]

[1]Nankai University

October 4, 2022

Abstract

The mutualistic lifestyle of pollinating fig wasps and fig trees provides an excellent model for studying ecological and adaptive evolution issues. Transposable elements (TEs), as an important component of the genomes, are powerful driver for organisms to adapt to environment. Here, the genomic TEs of six pollinating fig wasps and five non-pollinating fig wasps were analyzed in the characteristics of composition and their effects on genome size, the historical burst patterns and their association with effective population size and paleoclimate changes, to infer the role of TEs in environmental adaptation in fig wasps. Compared with non-pollinators, pollinators' TEs showed a significantly different burst state with less types and amount, shorter lengths, and lower contents in the genomes. The recent smaller effective population size and contractive demography failed to cause pollinators to accumulate more TEs, while the large number of TEs accumulated in non-pollinators positively correlated with their population expansion. The major TEs burst peaks in the history of pollinators highly overlapped with the warmer times in the Coolhouse in geological history. TEs located in the major peak period were mostly inserted near genes related to environment information processing such as Circadian entrainment pathway, and might act as CRMs (cis-regulatory modules) to regulate the conjunctive genes in response to paleoclimate changes in pollinators. These results revealed the molecular basis of the fig wasp's response to changes in the syconia microenvironment and paleoclimate macroenvironment from the perspective of genomic TEs.

Introduction

Fig pollinating wasps (pollinators) (Hymenoptera: Chalcidoidea), which symbiosis with fig trees (Moraceae, *Ficus*), is an excellent insect paradigm for studying ecological and adaptive evolution issues (Marussich and Machado 2007). Pollinators feed on the floret ovaries of the syconia of figs. The processes of embryonic development, maturation, mating, and female pollination and oviposition of pollinators are all completed inside the enclosed syconia, and the only step out of the syconia in their life histories is that the mated females fly out to look for other figs in the female phase. Once suitable figs are found, female pollinators immediately enter the syconia to lay eggs and pollinate the florets. In addition to pollinating fig wasps, a group of parasitic fig non-pollinating wasps (non-pollinators) also live in figs. Compared with pollinators, the life histories of non-pollinators related to figs are relatively complicated. According to diets, they can be classified into gallers, seed predators, kleptoparasites, and parasitoids (Borges 2015). Although the developmental process must be completed in the syconia, some winged males will fly out to mate, and the oviposition behaviors of females are mostly completed by penetrating through the syconium wall from the outside with their long ovipositors. Phylogenetic analyses suggested that the symbiosis of pollinators and non-pollinators with figs have evolved at least twice independently (Rasplus, Kerdelhue et al. 1998, Peters, Niehuis et al. 2018). Compared with non-pollinators, pollinators have dramatically reduced gene families involved in environmental sensing and detoxification, and have lost several genes in response to environmental stress and immune activation, many more genes were subject to relaxed selection (Xiao, Wei et al. 2021). Obviously, although both are related to figs, pollinators and non-pollinators have different life and evolutionary histories, and each of them has thus shown diverse adaptation characteristics in the aspect of functional gene.

1

The life history of fig wasps, especially pollinators, is closely connected with the enclosed fig fruits, which forms an inseparable relationship between the fates of wasps and their host fig trees. Fig trees are typical tropical and subtropical plants, and the distribution is expected to be strongly structured by climate for the sensitivity to geological climatic changes (Hawkins, Field et al. 2003). Molecular clock data has shown that the origin of figs is ca 75 million years ago (Ma) (Cruaud, Ronsted et al. 2012). Since the Cenozoic, the global climate has undergone a major and complex change (Zachos, Pagani et al. 2001). The most famous climate anomalies are the abnormally high temperature event at ~55 Ma, and the sudden temperature drop caused by the twice glaciation events at ~ 34 Ma (Ivany, Patterson et al. 2000) and ~23 Ma (Zachos, Shackleton et al. 2001). 34 Ma to 3.3 Ma is called The Coolhouse (Westerhold, Marwan et al. 2020). From 25 to 13.9 Ma, the reduction in ice volume allowed the Coolhouse to have a warmer times, called Miocene Climatic Optimum (~17 to 14 Ma) (Westerhold, Marwan et al. 2020). Rapid climate change is likely to cause a range of selection pressure on organisms, and evolutionary adaptation can be an important way for the latter to surmount rapid climate changes (Hoffmann and Sgro 2011). Since the responses of fig trees to paleoclimate changes have hardly been considered, due to the lack of good palynological records, we expect to find evidence that the evolution of these small insects closely related to the figs is affected by drastic paleoclimate anomalies, from the genomic perspective.

As an important component and source of variation in the genome, TEs are potent facilitators of genome rapid adaptive evolution (Legrand, Caron et al. 2019). The accumulation of TEs in the genome includes an increase in number and an expansion of types, and the most direct effect on the genome is the increase in genome size (Naville, Henriet et al. 2019). It has been reported that although there are large differences in the content of TEs in different species, there is generally positive correlation with genome size (Shao, Han et al. 2019). According to the "TE-Thrust" hypothesis, a suitable repertoire of TEs can increase the adaptability of species to environmental or ecological changes. Otherwise, the population may undergo a series of evolutionary phenomena, such as stasis, becoming "living fossils", or even becoming extinct (Oliver and Greene 2011). "Bursts" are the most rapid and direct way to increase the content of TEs, and the deeper reason is the adaptive evolution of the genome in response to rapid changes in the external environment induced by stress (Belyayev 2014, Schrader and Schmitz 2019). For example, the genome of the invasive ant *Cardiocondyla obscurior* possesses a quickly evolving accumulation of TEs (TE islands) distributed in the same region as the olfactory receptor gene, which is speculated to be related to the adaptive radiation (Schrader, Kim et al. 2014). The most common biological response to environmental adaptation is to regulate gene expression (Schrader and Schmitz 2019). For example, new insertion or excision of TEs can affect the transcription of genes, and TEs evolved into regulatory sequences, such as *cis* -regulatory modules (CRMs), which further affect gene expression (Oliver and Greene 2011, van't Hof, Campagne et al. 2016).

The different life and evolutionary histories of pollinators and non-pollinators imply that the two groups have different habitat adaptation processes. At the same time, the different life histories mean that they have substantially different effective population sizes (N_e), which may impact the process of TEs accumulation in the genomes of these species (Willemsen, Cui et al. 2020, Chak, Harris et al. 2021). Considering that the evolutionary mechanism of TEs in the process of environmental adaptation of fig wasps is still unclear, we first analyzed the composition and evolutionary differences of TEs in both groups of fig wasps, as well as the characteristics of burst patterns. We then explored the association of TEs burst with historical population dynamics and paleoclimate change and its molecular basis, aiming to uncover the adaptive changes of fig wasps to the syconia microenvironment and the geological climatic macroenvironment from the perspective of evolution of TEs.

Materials and Methods

TEs annotation and statistical analysis

Eleven fig wasps were selected for analysis, and details were listed in Table 1. The *de novo* repeat libraries were built using RepeatModeler (*http://www.repeatmasker.org/RepeatModeler/*, v2.0.2) and were filtered using species-specific protein annotation files. Then, homologous TEs library of 'chalcidoidea' was extracted using the perl script queryRepeatDarabase.pl. We then combined the filtered *de novo* re-

peat libraries and 'chalcidoidea' TEs library to obtain local species-specific TEs libraries. RepeatMasker (*http://www.repeatmasker.org*, v4.1.2) was used to mask the genomes. The parameters '-lib -s -nolow -norna -no_is -cutoff 250 -gff -align' were applied. To summarize the TEs annotation results, parseRM.pl (*https://github.com/4ureliek/Parsing-RepeatMasker-Outputs/*) script was used. Spearman correlation and Mann-Whitney U test were carried out using SPSS software.

Estimation of the TEs insertion time

The insertion time (T) was calculated using the adapted $T = K/2r$ formula (Li, Guo et al. 2019), where K is the sequence divergence, and r is the substitution rate. The formula for calculate sequence divergence is K=-300/4*Ln(1-D*4/300), where D is converted from the mutation rate d (divergence) of each site in the file with suffix of 'align' as one of the RepeatMasker results. The formula is D=d/100.

To estimate the nucleotide substitution rate r , 4-fold degenerate sites were extracted from single-copy orthologs of 17 species in our previous report (Xiao, Wei et al. 2021). Further, taking the known tree topology as input parameters, we used the phyloFit program in the PHAST (*http://compgen.cshl.edu/phast/*) package to estimate the branch length of the input evolutionary tree. The root-to-tip branch length was calculated by the TreeStat tool (*http://tree.bio.ed.ac.uk/software/treestat/*). The substitution rate was calculated using the root-to-tip branch length divided by the divergence time (485 Mya) of *D. pulex* on the phylogenetic tree in the report (Xiao, Wei et al. 2021).

PSMC analysis

Illumina short reads were aligned to the fig wasp genomes with BWA-MEM (Li and Durbin 2009). The consensus sequences were called by SAMtools and BCFtools (Li 2011). The argument -C50 was passed to BCFtools mpileup to adjust mapping qualities, and variants were called using BCFtools call with parameter -c and then piped to vcf2fq to generate consensus FASTQ files with the parameter -d 30 and -D 200. Next, FASTQ files were converted to FASTA-like psmafa file using fq2psmcfa with the parameter -q 20. PSMC (Li and Durbin 2011) was run with parameters -N25 -t15 -r 5 -p "4+25*2+4+6" for coalescent simulation. The neutral mutation rate per generation per site (μ) were calculated from the nucleotide substitution rate per million year per site (r) and is listed in Table 1.

Function enrichment analysis on genes near TEs

The gene sets within 1 Kb near the classified TEs that were inserted during the periods of major peaks were selected. For Pcor, Ekon, Wpum, Dvas, Kgib, and Cfus, the time intervals were 24-26 Ma, 20-21 Ma, 20-22 Ma, 11-12 Ma, 20-21 Ma, 15-16 Ma in turn. GO and KEGG enrichment analysis were carried out using the plug-in ClueGO (Bindea, Mlecnik et al. 2009) in Cytoscape software with default parameters. Based on the genes shared between terms, term-term interrelation was defined using kappa score. To clearly observe the commonalities, the enrichment networks were merged in the intersection mode. Then, the nodes were filtered using the parameter Term P- Value < 0.05.

cis-regulatory modules (CRMs) prediction

To avoid the potential 'false' high score to regions with one or more occurrences of the repeated pattern, Tandem Repeats Finder (TRF, v4.09) (Benson 1999) was used to masked tandem repeats in genome. Then SCRMshaw (Kazemian and Halfon 2019) was run with the *Drosophila* training data, and with parameters '–thitw 2000 –wlen 200 –wshift 100 –gff –imm –hexmcd –pac –genome –traindirlst –outdir' in the masked genomes. Finally, the needed hits were chosen using Generate_top_N_SCRMhits.pl script, and sorted and merged by BEDTools (Quinlan and Hall 2010). The following comparing SCRMshaw predictions to TEs was carried out by using BEDTools.

Results

TEs composition of fig wasps and its effect on genome size

In this study, TEs were annotated and analyzed in the genomes of six pollinators *Platyscapa corneri* (Pcor),

3

Eupristina koningsbergeri (Ekon), *Wiebesia pumilae* (Wpum), *Dolichoris vasculosae* (Dvas), *Kradibia gibbosae* (Kgib), *Ceratosolen fusciceps* (Cfus) and five non-pollinators *Sycobia* sp.2 (Sbsp),*Philotrypesis tridentata* (Ptri), *Apocrypta bakeri* (Abak), *Sycophaga agraensis* (Sagr), *Sycophila* sp.2 (Spsp) (Table 1). In terms of TEs types, the classified TEs families in pollinators varied from 18 to 31, while in non-pollinators from 35 to 52. Among the pollinators, Dvas had the smallest number and the shortest total length of TEs, with 17,660 totaling 5.78 Mb. Ekon had the largest number and the longest total length of TEs, with 104,360 totaling 47.61 Mb. The non-pollinator Sagr had the smallest number of TEs (109,977) which was still more than pollinator Ekon. Abak had the shortest total length of 25.17 Mb. Sbsp had the largest number and the longest total length of TEs, up to 647,195 totaling 231.12 Mb. Clearly, there were significantly fewer families and quantity, and shorter total length of TEs in pollinators than that in non-pollinators.

To clarify the effect of TEs on genome size, the Spearman correlation coefficient between the total length of TEs and genome size in fig wasps was analyzed and the result was significantly positive ($R^2 = 0.618$, $P = 0.043$) (Figure 1B). Further, the differences in TEs content and genome size between the two groups of fig wasps were analyzed. The TEs percentage varied from 2.02% to 15.12% in pollinators, and ranged from 12.69% to 41.71% in non-pollinators (Table 1). The Mann-Whitney U test showed that the TEs content of pollinators was significantly lower than that of non-pollinators ($P = 0.017$) (Figure 1C), but the difference in genome size between the two groups was not significant. A possible reason is that the two groups of species involved are not reciprocally monophyletic (Xiao, Wei et al. 2021). Even so, the species Sbsp with the longest total length of TEs and Ptri with the highest TEs content, were also had the largest and second-largest genome sizes, respectively. Obviously, TEs directly drive the genome size expansion.

To elucidate the specific ways in which TEs affect genome size, the genome proportion of each classified TEs family in each species were calculated. The results showed that more than 50% of the classified TEs contents were derived from one to four types of TEs families (Table S1), and the TEs family with the highest content in each species was not the same. For pollinators, the TEs family with the highest content in Pcor, Dvas, Kgib, and Cfus was RTE-X, I-Jockey, R1-LOA, and R1, with a size of 4.96 Mb, 2.18 Mb, 0.92 Mb, and 1.07 Mb, respectively. In Ekon and Wpum, it was Gypsy with a size of 7.71 Mb and 1.90 Mb, respectively. For non-pollinators, the most abundant TEs family were all Gypsy, and the sizes varied between 2.28 to 46.05 Mb. Therefore, a few types of TEs bursts caused a large accumulation of TEs in the genome and thus affected the genome size.

Burst events of TEs retained in each species

In order to analyze the characteristics of the burst patterns caused by its differences in the composition of TEs, we drew stacked plots of the total TEs length within the genome of each species by insertion time, and obtained the distribution map of TEs insertion time (Figure 2). The burst peaks accumulated by different types of TEs in a specific time could be found in each species, and we defined that the highest peak was the major burst peak, and other peaks were minor peaks. It was easy to find that the burst patterns between the two groups of fig wasps were inconsistent. The burst peaks of TEs in pollinators showed high in the middle and low on both sides. Among them, Pcor, Ekon, Kgib and Cfus had only the single major peaks, no obvious minor peak, which looks like "n". While Wpum and Dvas had obvious minor peaks in addition to the major peaks, which is shaped like "m". In most non-pollinators, the TEs burst peaks presented a continuous distribution of high left and low right. For examples, the distribution of TEs in Sbsp, Ptri, Sagr, and Spsp were like a slide, which was called "ladder" type. Only the TEs distribution in Abak resembled the "m" type. To eliminate the interference caused by differences in fig trees, the TEs distribution patterns of a pollinating fig wasp Ekon and three non-pollinating fig wasps Sbsp, Ptri, and Spsp co-living on*Ficus benjamina* were compared. The "n"-shaped peak of pollinator contrasted sharply with the "ladder"-shaped distribution of non-pollinators.

Furthermore, the evolutionary history of TE bursts was detected based on the time axis. The peaks for pollinators were concentrated and more ancient, while the peaks for non-pollinators were scattered and dominated by recent events. The historical time of the major burst peak of TEs was calibrated for each species. The major burst peak of pollinators was in the period of 11-24 Ma, while it was in 1-17 Ma of

4

non-pollinators. In evolutionary history, the major burst peak in a species is implied to be the largest TE component change reserved in the genome related to environmental adaptation. Mapping the TEs burst time with the paleoclimate history, it was not difficult to find that the warmer times in the Coolhouse period in geological history coincided with the distribution time of the major burst peak of pollinators. Based on this, we suggested that the ancient traces of TEs burst preserved in the genome of pollinating fig wasps were closely related to the adaptation to paleoclimate changes.

Historical population dynamics of fig wasps

It has been reported that more genes in pollinating fig wasps were subjected to positive selection compared with non-pollinators, probably due to the chronically extremely small population size of pollinators (Xiao, Wei et al. 2021). Extremely small population size may lead to the accumulation of deleterious mutations like TEs (Chak, Harris et al. 2021). To analyze the relationship between the burst history of TEs in the fig wasp genomes and the historical population dynamics, we used the pairwise sequentially Markovian coalescent (PSMC) model to tract historical fluctuations in species' effective population size (N_e). As shown in Figure 3, from approximately 1 Ma to 10 thousand years ago (Kya), the N_e curves of pollinators fluctuated and shrinked in different ways. Pcor and Dvas showed gradual decline in N_e. The N_e of Dvas and Cfus increased first and then decreased below the starting point. The front part of the curve of Wpum was similar to that of Dvas, but then it picked up a little, and also did not exceed the starting point. In Kgib, there was an initial increased in N_e followed by two valley-peak alternations, however, the results were inaccurate in the last 10 Kya, possibly due to recombination events (Li and Durbin 2011). In non-pollinators, the N_e trajectories were not uniform, and dominated by fluctuation and expansion from ~1 Ma to ~10 Kya. The N_e had an initial gradual increase in Sbsp and Spsp, then went through 3 peaks and finally was higher than the initial point. The N_e of Abak showed a simple continuous rise. After experiencing valley-peak alternations, Sagr showed no significant difference in N_e, but the N_e of Ptri was significantly lower than the beginning. In contrast, the N_e of pollinators contracted and was smaller than that of non-pollinators. However, within 1 Ma, the accumulation of TEs was rare in pollinating fig wasps, while it continued to increase in non-pollinating fig wasps. Therefore, pollinators did not retain more TEs due to the smaller N_e.

Gene functions near the burst TEs of pollinating fig wasps

To clarify possible historical environmental factors for TEs burst in fig pollinators, we analyzed the functions of genes near TEs based on the fact that the insertion or deletion of TEs may have a strong impact on the expression of nearby genes. Enrichment analysis was performed for genes within 1 Kb near the classified TEs during the period of the major burst peaks of each pollinator, then took the intersection of all enrichment results. As shown in Figure 4, the results of GO enrichment analysis indicated that genes related to ion transport and signal transduction processes were significantly enriched, while KEGG enrichment analysis showed that the Circadian entrainment pathway (map 04713), and nervous system and signal related pathway were significantly enriched. That is, TEs in the major burst peaks of pollinating fig wasps were inserted in the vicinity of genes related to environmental information processing.

TEs act as CRMs affecting gene expression in fig pollinators

To elucidate the mechanism by which TEs affect the environmental adaptation of fig pollinators, the evolutionary association between TEs and *cis*-regulatory modules (CRMs) was analyzed based on the prevalence of TEs evolving into regulatory sequences. First, the genome-wide CRMs were predicted, and then the location information was matched between CRMs and TEs. As a result, 6.6%-23.5% of TEs highly overlapped with CRMs (Table S2). Among them, Kgib accounted for the highest proportion.

Taking Kgib as an example, the genomic distribution of TEs and CRMs was plotted, and the results showed that the density distributions of CRMs and TEs were just opposite (Figure 5). That is, the location with more CRMs had less TEs, and the location with more TEs had less CRMs. This may be due to the fact that CRMs are closer to genes (Blanchette, Bataille et al. 2006), but TEs near genes are more likely to be deleted (Sultana, Zamborlini et al. 2017). However, there were still a lot of overlaps between TEs and CRMs in three modes (Figure 5). Among them, the mode of TEs located in CRMs accounted for the largest

5

proportion, close to 50%, followed by the intersection of TEs and CRMs. The least proportion is the mode of CRMs in TEs, only 12.2%. That is to say, in a few cases, the CRMs fragments in the TEs sequence were retained, and more often, TEs acted as CRMs to affect gene expression.

KEGG enrichment analysis was also performed on the gene near the CRMs that overlapped with TEs in the major burst peaks. The results showed that the pathways involved in environment information processing, such as Circadian entrainment, were also enriched in most species (Figure S1).

Discussion

Fig wasps are a group of special insects that live in the enclosed syconia of fig trees. In the long evolutionary history, they have maintained a close ecological relationship with the fig fruits. Related to the shelter and relatively safe living environment provided by fig fruits, the evolution of the fig wasps shows an adaption to the environment within fig syconia, however, their adaption to the large-scale paleoclimatic environment may also be closely related to the life in the fig syconia. This study focused on TEs in the genomes and tried to explore the evolutionary traces of the fig wasps in response to the changes of the microenvironment of fig syconia and the paleoclimatic macroenvironment.

Factors influencing the composition and evolution of TEs in fig wasps

We compared the differences in TEs composition between six pollinator and five non-pollinator species. The results displayed that the types, quantity and total length of TEs in pollinators were significantly less than those in non-pollinators. This may be linked to the divergent life histories between the pollinators and the non-pollinators. In pollinator species, the stricter host specificity, single and sufficient food sources, stable and relatively hidden habitats, may be associated with insufficient motivation for expansion of TEs and thus the lack of TEs in the genomes (Machado, Robbins et al. 2005, Cruaud, Ronsted et al. 2012). In contrast, the non-pollinators can usually live on several host fig species, with various parasitic strategies and complex feeding habits, which requires a stronger ability to detect the hosts and foods (Borges 2015). Besides, the longer time of non-pollinators living in the external surface of syconia for mating and oviposition is fraught with the risks of predation by ants and diseases infection (Ranganathan, Ghara et al. 2010), and thus also requires a stronger ability to avoid natural enemies and resist diseases. In addition, the highly inbreeding reproduction of pollinating fig wasps greatly reduces the genetic diversity of the species (Greeff, Jansen van Vuuren et al. 2009), while winged males of non-pollinators can fly out of their native fig fruits and mate outside, which will reduce the probability of inbreeding (Cook, Reuter et al. 2017). All these characters provide a greater possibility for non-pollinators to increase the types of TEs, especially the number and total length.

Inbreeding reproduction will influence the mutation accumulation in genomes by affecting the N_e (Lefebure, Morvan et al. 2017, Bonchev and Willi 2018), as small N_e will exacerbate accumulation of deleterious mutations including TEs (Willemsen, Cui et al. 2020). Within 1 Ma, the N_e of pollinating fig wasps was smaller and in a state of fluctuating contraction, which was consistent with the situation that many genes in this group were subjected to positive selection. In contrast, the N_e of non-pollinating fig wasps was larger and dominated by dynamic expansion. However, for the distribution of TEs, the accumulation of TEs did not show a strong burst trend in pollinators, while in non-pollinators the accumulation of TEs continued to increase. This is inconsistent with the reported theory (Kimura 1983) that smaller N_e favor mutation accumulation (Chak, Harris et al. 2021). This may be due to the protective effect of syconia, coupled with a highly inbreeding reproductive system, which has a far greater impact on the accumulation of TEs than the N_e in pollinators. Because of the relatively complex habitat of non-pollinators and the lower probability of inbreeding, the genetic diversity of species increases. This leads to a continuous burst of TEs, which improves the fitness of the population, and ultimately make the N_e showed an expansion state. There were also exceptions. For example, the contractive N_e of Ptri may be related to the intense competition from other non-pollinators living on the same host (Limberger and Fussmann 2021).

Next, the differences in the burst patterns further illustrate the different habitat adaptation processes of both groups of fig wasps. Taking Gypsy as an example, their expansion were widespread in fig wasp species,

6

so the contents of TEs in the genomes of both pollinators and non-pollinators had increased significantly. However, the patterns of the Gypsy burst in both groups were quite different, with the degree of expansion higher in non-pollinators, so there was a significant difference in the TEs contents in the genomes of both groups. In order to more clearly reflect the differences in the composition of TEs in evolutionary history, the insertion time distribution of TEs were drawn. Comparing the burst patterns of both groups, we found that the TEs of pollinators showed a state of high in the middle and low on both sides, while non-pollinators showed a continuous distribution of high on the left and low on the right. This distribution mode was contrasted sharply between both groups of species living on *Ficus benjamina* . The protective effect of figs on pollinators may make them sluggish or insensitive to the external environment changes, which may lead to the burst of TEs occurring only in extreme environments. The evolutionary pressure of complex external habitats on non-pollinators persisted, which is consistent with the "ladder" type of TEs increases in this group.

The association of TEs burst and paleoclimate change and its molecular basis

Our study detected that the pollinating fig wasps had significant TEs burst events in the time range of 11-24 Ma. This period coincided with the warmer times during the Coolhouse period in geological history (Westerhold, Marwan et al. 2020). Due to the action of the *Oi-1* glaciation at the Eocene-Oligocene transition period, the global temperature dropped significantly and the Antarctic ice sheet grew rapidly (Zachos, Pagani et al. 2001). In the late Oligocene (~25 Ma), a warming current reduced the extent of Antarctic ice sheet. This warm climate lasted until the mid-Miocene, and reached its peak at ~13.9 Ma. Thereafter, the East Antarctic Ice Sheet(EAIS)began to expand, and the global climate underwent a major reorganization (Groeneveld, Henderiks et al. 2017). It has been reported that the emergence of fig wasps is distinctly seasonal. During the rainy season, continuous rainy day caused a large number of pollinators to die after emergence, and the high temperature also shortened the lifespan of them, resulting in a low reproductive success rate. While in the dry season, although emergence rate is reduced, the dry climate is beneficial for flight, and the cool temperature prolongs the lifespan of pollinators, thereby significantly increasing reproductive success rate (Wang, Yang et al. 2005, Dong, Peng et al. 2020). Considering the distribution of fig trees in the tropical and subtropical regions, it can be seen that the low temperature within a certain range is more suitable for the reproduction of pollinators. That is, the high temperature period during the Coolhouse as strong selective pressure may be a historical challenge for pollinator fig wasps. Non-pollinators are less sensitive to climate change, and their survial pressures come from competition with other fig wasps (Weiblen 2002, Zhang and Yang 2009, Wang, Compton et al. 2015).

To further clarify the relationship between the TEs bursts in pollinating fig wasps and the paleoclimate changes, genes near TEs that were inserted during the periods of major peaks were analyzed. KEGG analysis confirmed that Circadian entrainment pathway related genes involved in environmental adaptation were significantly enriched. Circadian entrainment is the process to activate the internal biological clock of the organism by recurring exogenous signals, so that the organism's endocrine and behavioral rhythms will be synchronized to environmental cues (Johnson, Elliott et al. 2003). It is an interesting and open question whether the periodic climate changes will have a direct impact on the fig wasps, or they will exert selective pressure on the fig wasps through their influence on the fig trees. Generally, climate changes have great impacts on plants because they can change the flowering cycle (Ehrlen and Valdes 2020). The year-round phenological variations of fig trees living in the tropics are generally limited, with fig production featuring intra-individual flowering synchrony and inter-individual asynchrony throughout the year (Bronstein, Gouyon et al. 1990). However, for the higher latitude regions and the edges of the species' distribution range, the synchronization of flowering synchrony, and the productions of wasps and seeds display seasonal fluctuations, which are accompanied by a reduction in reproductive success (Chen, Zhang et al. 2018). Therefore, considering that both parts of the symbiosis system of the fig and fig pollinators have different responses to the geographic climate differences, it is not difficult to speculate that, in the evolutionary history of the symbiosis system, the changes in the flowering periods induced by the rapid paleoclimate changes would exert great evolutionary pressure on the pollinators, because they need to keep pace with the flowering cycle of the fig trees. In this adaptive process, the adaptation of the Circadian entrainment genes may play a

pivotal role. In other words, periodic climatic changes can have a huge impact on plants, which in turn can affect the evolution of symbiotic organisms closely related to plants.

Finally, by analyzing the correlation between the distributions of TEs and CRMs in the genomes of the fig pollinators, we speculated a potential molecular mechanism for the adaptive evolution of these wasps through the action of TEs. There are three main ways for TEs to regulate host genes. One is that newly inserted or deleted TEs may change the regulation or coding environment at a specific location, thereby affecting gene expression. For example, the adaptive "industrial melanism" phenotype of *Biston betularia* has been evolved through the intronic insertion of a DNA transposon, increasing transcript abundance of the *cortex* gene during early wing development (van't Hof, Campagne et al. 2016). Second, TEs insertions might also lead to modifications at the epigenetic level, which can affect the expression of neighboring genes. For example, in*Arabidopsis* , gene expressions are negatively correlated with the density of methylated TEs (Hollister and Gaut 2009). Third and the last, TEs could be evolved into regulatory sequences which might affect the downstream gene expression (Oliver and Greene 2011). In the view of the positive role and universality of retained TEs as regulatory sequences, we analyzed the relationship between TEs and CRMs. The results are consistent with previous researches that there are many TEs overlapping with CRMs. Most of the overlapped TEs are in the modes of intersection or in CRMs. That is to say, many TEs in the genome may act as CRMs to closely regulate gene expression to achieve environmental adaption. The genes near those CRMs sequences that overlapped with TEs also included the genes in the pathway related to environmental information processing such as Circadian entrainment. This suggested that bursts TEs in these species may act as CRMs to influence nearby gene expression to achieve gene environmental adaptability.

Conclusion

The fig wasps live in the enclosed syconia of fig trees that are restricted in tropical and subtropical regions. TEs are one of the powerful drivers of the adaptive evolution of organisms. The composition and burst patterns of the genomic TEs in fig wasps reflect their adaptation to the syconia microenvironment. Meanwhile, from the perspective of adaptation to the global climate changes, when the fig trees affected by geological climatic changes and underwent drastic changes in the physiological characters such as the flowering cycle, the pollinators may self-regulate to adapt to these changes in its habitat of the syconia by regulation of genes related to Circadian entrainment pathways through the changes of TEs in the genomes. In the future, further verification of the functions of TEs on the regulation of related genes will help us understand how this closely related symbiotic system between the figs and fig wasps has been adapted to environmental changes during the evolutionary process.

References

Belyayev, A. (2014). "Bursts of transposable elements as an evolutionary driving force." Journal of Evolutionary Biology 27(12): 2573-2584.

Benson, G. (1999). "Tandem repeats finder: a program to analyze DNA sequences." Nucleic Acids Research 27(2): 573-580.

Bindea, G., B. Mlecnik, H. Hackl, P. Charoentong, M. Tosolini, A. Kirilovsky, W. H. Fridman, F. Pages, Z. Trajanoski and J. Galon (2009). "ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks." Bioinformatics 25(8): 1091-1093.

Blanchette, M., A. R. Bataille, X. Y. Chen, C. Poitras, J. Laganiere, C. Lefebvre, G. Deblois, V. Giguere, V. Ferretti, D. Bergeron, B. Coulombe and F. O. Robert (2006). "Genome-wide computational prediction of transcriptional regulatory modules reveals new insights into human gene expression." Genome Research 16(5): 656-668.

Bonchev, G. and Y. Willi (2018). "Accumulation of transposable elements in selfing populations of *Arabidopsis lyrata* supports the ectopic recombination model of transposon evolution." New Phytol 219(2): 767-778.

Borges, R. M. (2015). "How to be a fig wasp parasite on the fig-fig wasp mutualism." Current Opinion in Insect Science 8: 34-40.

Bronstein, J. L., P. H. Gouyon, C. Gliddon, F. Kjellberg and G. Michaloud (1990). "The Ecological Consequences of Flowering Asynchrony in Monoecious Figs: A Simulation Study." Ecology 71(6): 2145-2156.

Chak, S. T. C., S. E. Harris, K. M. Hultgren, N. W. Jeffery and D. R. Rubenstein (2021). "Eusociality in snapping shrimps is associated with larger genomes and an accumulation of transposable elements." Proc Natl Acad Sci U S A 118(24).

Chen, H., Y. Zhang, Y. Peng and R. T. Corlett (2018). "Latitudinal effects on phenology near the northern limit of figs in China." Scientific Reports 8(1): 4320.

Cook, J. M., C. Reuter, J. C. Moore and S. A. West (2017). "Molecular markers reveal reproductive strategies of non-pollinating fig wasps." Ecological Entomology 42(6): 689-696.

Cruaud, A., N. Ronsted, B. Chantarasuwan, L. S. Chou, W. L. Clement, A. Couloux, B. Cousins, G. Genson, R. D. Harrison, P. E. Hanson, M. Hossaert-McKey, R. Jabbour-Zahab, E. Jousselin, C. Kerdelhue, F. Kjellberg, C. Lopez-Vaamonde, J. Peebles, Y. Q. Peng, R. A. Pereira, T. Schramm, R. Ubaidillah, S. van Noort, G. D. Weiblen, D. R. Yang, A. Yodpinyanee, R. Libeskind-Hadas, J. M. Cook, J. Y. Rasplus and V. Savolainen (2012). "An extreme case of plant-insect codiversification: figs and fig-pollinating wasps." Syst Biol 61(6): 1029-1047.

Dong, Y., Y. Peng and B. Wang (2020). "Seasonal dynamics of fig wasp community and interaction networks in Ficus benjamina." Biodiversity Science 28(4): 496-503.

Ehrlen, J. and A. Valdes (2020). "Climate drives among-year variation in natural selection on flowering time." Ecology Letters 23(4): 653-662.

Greeff, J. M., G. J. Jansen van Vuuren, P. Kryger and J. C. Moore (2009). "Outbreeding and possibly inbreeding depression in a pollinating fig wasp with a mixed mating system." Heredity (Edinb) 102(4): 349-356.

Groeneveld, J., J. Henderiks, W. Renema, C. M. McHugh, D. De Vleeschouwer, B. A. Christensen, C. S. Fulthorpe, L. Reuning, S. J. Gallagher, K. Bogus, G. Auer, T. Ishiwa and S. Expedition (2017). "Australian shelf sediments reveal shifts in Miocene Southern Hemisphere westerlies." Science Advances 3(5): e1602567.

Hawkins, B. A., R. Field, H. V. Cornell, D. J. Currie, J. F. Guegan, D. M. Kaufman, J. T. Kerr, G. G. Mittelbach, T. Oberdorff, E. M. O'Brien, E. E. Porter and J. R. G. Turner (2003). "Energy, water, and broad-scale geographic patterns of species richness." Ecology 84(12): 3105-3117.

Hoffmann, A. A. and C. M. Sgro (2011). "Climate change and evolutionary adaptation." Nature 470(7335): 479-485.

Hollister, J. D. and B. S. Gaut (2009). "Epigenetic silencing of transposable elements: A trade-off between reduced transposition and deleterious effects on neighboring gene expression." Genome Research 19(8): 1419-1428.

Ivany, L. C., W. P. Patterson and K. C. Lohmann (2000). "Cooler winters as a possible cause of mass extinctions at the Eocene/Oligocene boundary." Nature 407: 887-890.

Johnson, C. H., J. A. Elliott and R. Foster (2003). "Entrainment of circadian programs." Chronobiology International 20(5): 741-774.

Kazemian, M. and M. S. Halfon (2019). "CRM Discovery Beyond Model Insects." Methods Mol Biol 1858: 117-139.

Kimura, M. (1983). The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge University Press.

Lefebure, T., C. Morvan, F. Malard, C. Francois, L. Konecny-Dupre, L. Gueguen, M. Weiss-Gayet, A. Seguin-Orlando, L. Ermini, C. Sarkissian, N. P. Charrier, D. Eme, F. Mermillod-Blondin, L. Duret, C. Vieira, L. Orlando and C. J. Douady (2017). "Less effective selection leads to larger genomes." Genome Res 27(6): 1016-1028.

Legrand, S., T. Caron, F. Maumus, S. Schvartzman, L. Quadrana, E. Durand, S. Gallina, M. Pauwels, C. Mazoyer, L. Huyghe, V. Colot, M. Hanikenne and V. Castric (2019). "Differential retention of transposable element-derived sequences in outcrossing *Arabidopsis* genomes." Mob DNA 10.

Li, H. (2011). "A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data." Bioinformatics 27(21): 2987-2993.

Li, H. and R. Durbin (2009). "Fast and accurate short read alignment with Burrows-Wheeler transform." Bioinformatics 25(14): 1754-1760.

Li, H. and R. Durbin (2011). "Inference of human population history from individual whole-genome sequences." Nature 475(7357): 493-496.

Li, S. F., Y. J. Guo, J. R. Li, D. X. Zhang, B. X. Wang, N. Li, C. L. Deng and W. J. Gao (2019). "The landscape of transposable elements and satellite DNAs in the genome of a dioecious plant spinach (*Spinacia oleracea* L.)." Mob DNA 10: 3.

Limberger, R. and G. F. Fussmann (2021). "Adaptation and competition in deteriorating environments." Proc Biol Sci 288(1946): 20202967.

Machado, C. A., N. Robbins, M. T. Gilbert and E. A. Herre (2005). "Critical review of host specificity and its coevolutionary implications in the fig/fig-wasp mutualism." Proc Natl Acad Sci U S A 102 Suppl 1: 6558-6565.

Marussich, W. A. and C. A. Machado (2007). "Host-specificity and coevolution among pollinating and nonpollinating New World fig wasps." Molecular Ecology 16(9): 1925-1946.

Naville, M., S. Henriet, I. Warren, S. Sumic, M. Reeve, J. N. Volff and D. Chourrout (2019). "Massive Changes of Genome Size Driven by Expansions of Non-autonomous Transposable Elements." Curr Biol 29(7): 1161-1168 e1166.

Oliver, K. R. and W. K. Greene (2011). "Mobile DNA and the TE-Thrust hypothesis: supporting evidence from the primates." Mob DNA 2(1): 8.

Peters, R. S., O. Niehuis, S. Gunkel, M. Blaser, C. Mayer, L. Podsiadlowski, A. Kozlov, A. Donath, S. van Noort, S. Liu, X. Zhou, B. Misof, J. Heraty and L. Krogmann (2018). "Transcriptome sequence-based phylogeny of chalcidoid wasps (Hymenoptera: Chalcidoidea) reveals a history of rapid radiations, convergence, and evolutionary success." Mol Phylogenet Evol 120: 286-296.

Quinlan, A. R. and I. M. Hall (2010). "BEDTools: a flexible suite of utilities for comparing genomic features." Bioinformatics 26(6): 841-842.

Ranganathan, Y., M. Ghara and R. M. Borges (2010). "Temporal associations in fig-wasp-ant interactions: diel and phenological patterns." Entomologia Experimentalis Et Applicata 137(1): 50-61.

Rasplus, J. Y., C. Kerdelhue, I. Le Clainche and G. Mondor (1998). "Molecular phylogeny of fig wasps Agaonidae are not monophyletic." Comptes Rendus De L Academie Des Sciences Serie Iii-Sciences De La Vie-Life Sciences 321(6): 517-527.

Schrader, L., J. W. Kim, D. Ence, A. Zimin, A. Klein, K. Wyschetzki, T. Weichselgartner, C. Kemena, J. Stokl, E. Schultner, Y. Wurm, C. D. Smith, M. Yandell, J. Heinze, J. Gadau and J. Oettler (2014). "Transposable element islands facilitate adaptation to novel environments in an invasive species." Nat Commun 5: 5495.

Schrader, L. and J. Schmitz (2019). "The impact of transposable elements in adaptive evolution." Molecular Ecology 28(6): 1537-1549.

Shao, F., M. J. Han and Z. G. Peng (2019). "Evolution and diversity of transposable elements in fish genomes." Scientific Reports 9.

Sultana, T., A. Zamborlini, G. Cristofari and P. Lesage (2017). "Integration site selection by retroviruses and transposable elements in eukaryotes." Nat Rev Genet 18(5): 292-308.

van't Hof, A. E., P. Campagne, D. J. Rigden, C. J. Yung, J. Lingley, M. A. Quail, N. Hall, A. C. Darby and I. J. Saccheri (2016). "The industrial melanism mutation in British peppered moths is a transposable element." Nature 534(7605): 102-105.

Wang, R., S. G. Compton, R. J. Quinnell, Y. Q. Peng, L. Barwell and Y. Chen (2015). "Insect responses to host plant provision beyond natural boundaries: latitudinal and altitudinal variation in a Chinese fig wasp community." Ecol Evol 5(17): 3642-3656.

Wang, R. W., J. X. Yang and D. R. Yang (2005). "Seasonal Changes in the Trade-off Among Fig-supported Wasps andViable Seeds in Figs and Their Evolutionary Implications." J Integr Plant Biol 47(2): 144-152.

Weiblen, G. D. (2002). "How to be a fig wasp." Annual Review of Entomology 47: 299-330.

Westerhold, T., N. Marwan, A. J. Drury, D. Liebrand, C. Agnini, E. Anagnostou, J. S. K. Barnet, S. M. Bohaty, D. De Vleeschouwer, F. Florindo, T. Frederichs, D. A. Hodell, A. E. Holbourn, D. Kroon, V. Lauretano, K. Littler, L. J. Lourens, M. Lyle, H. Palike, U. Rohl, J. Tian, R. H. Wilkens, P. A. Wilson and J. C. Zachos (2020). "An astronomically dated record of Earth's climate and its predictability over the last 66 million years." Science 369(6509): 1383-1387.

Willemsen, D., R. Cui, M. Reichard and D. R. Valenzano (2020). "Intra-species differences in population size shape life history and genome evolution." Elife 9.

Xiao, J. H., X. Q. Wei, Y. Zhou, Z. Z. Xin, Y. H. Miao, H. X. Hou, J. X. Li, D. Zhao, J. Liu, R. Chen, L. M. Niu, G. C. Ma, W. Q. Zhen, S. M. He, J. X. Wang, X. F. Wei, W. H. Dou, Z. X. Sui, H. K. Zhang, S. L. Xing, M. Shi and D. W. Huang (2021). "Genomes of 12 fig wasps provide insights into the adaptation of pollinators to fig syconia." Journal of Genetics and Genomics 48(3): 225-236.

Zachos, J., M. Pagani, L. Sloan, E. Thomas and K. Billups (2001). "Trends, rhythms, and aberrations in global climate 65 Ma to present." Science 292(5517): 686-693.

Zachos, J. C., N. J. Shackleton, J. S. Revenaugh, H. Palike and B. P. Flower (2001). "Climate response to orbital forcing across the Oligocene-Miocene boundary." Science 292(5515): 274-278.

Zhang, F. P. and D. R. Yang (2009). "Oviposition timing and community structure of Ficus curtipes fig wasps." Chinese Journal of Applied Ecology 20(8): 2005-2011.

**Data Accessibility Statement**

Raw Illumina sequencing data will be deposited in the SRA. Genomic data used for TEs annotation are available from NCBI (reference = https://doi.org/10.1016/j.jgg.2021.02.010).

Figures and Tables
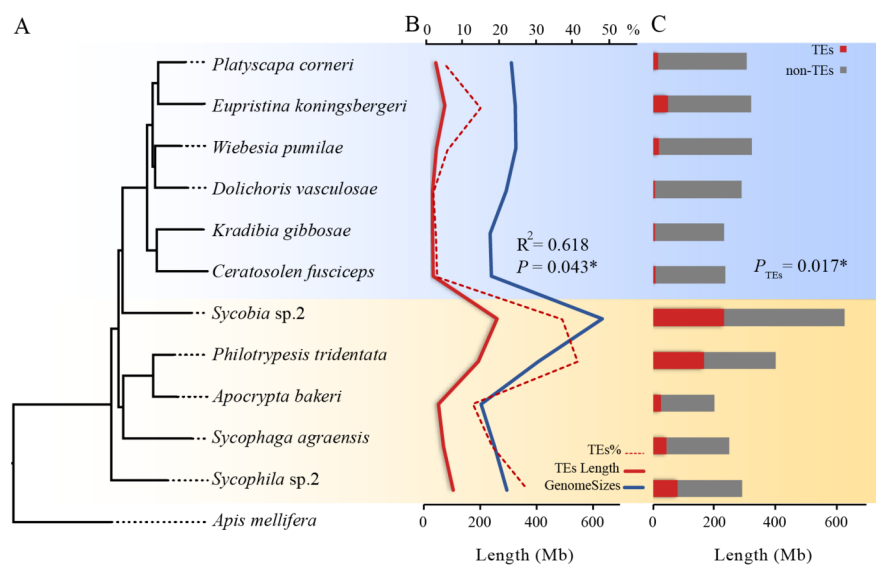
Figure 1: Summary of fig wasps genomic TEs contents. Purple and yellow background highlight pollinators and non-pollinators, respectively. (A) Phylogeny of samples in our study. (B) Correlation between TEs contents and genome sizes was significantly positive (Spearman correlation). (C) Summarizes the TEs and non-TEs length in the genomes, with significant difference in TEs length (Mann-Whitney U test) but not in genome sizes between pollinators and non-pollinators. These asterisks indicate a significant difference between pollinators and non-pollinators ($P <0.05$).
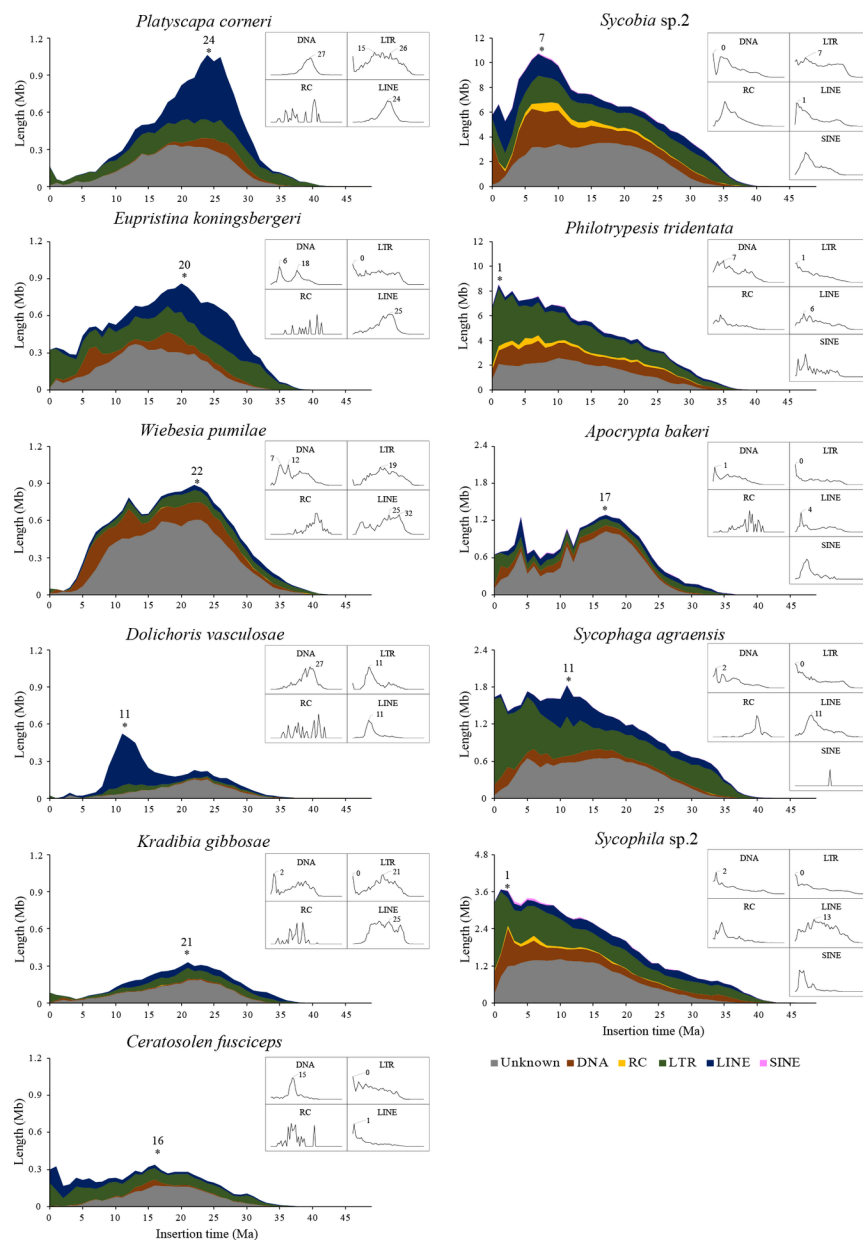
Figure 2: Distributiondiagrams of the TEs insertion time. In the diagram of each species, values of x-axis represent insertion time in million years ago (Ma), and y-axis represent TEs length in million base pairs (Mb). Peaks represent bursts of TEs insertion in the genome. Asterisks indicate the major peaks in each species.
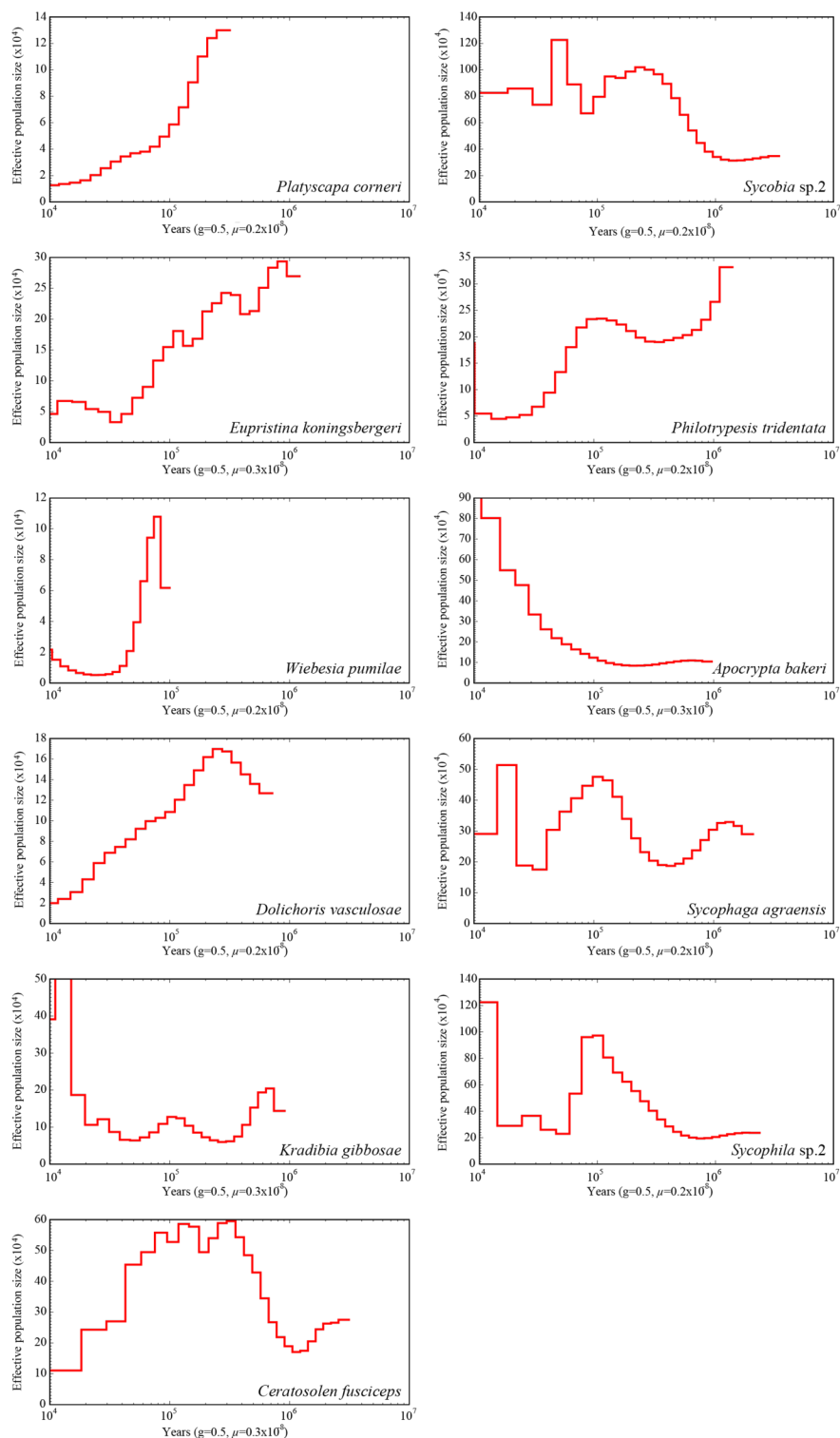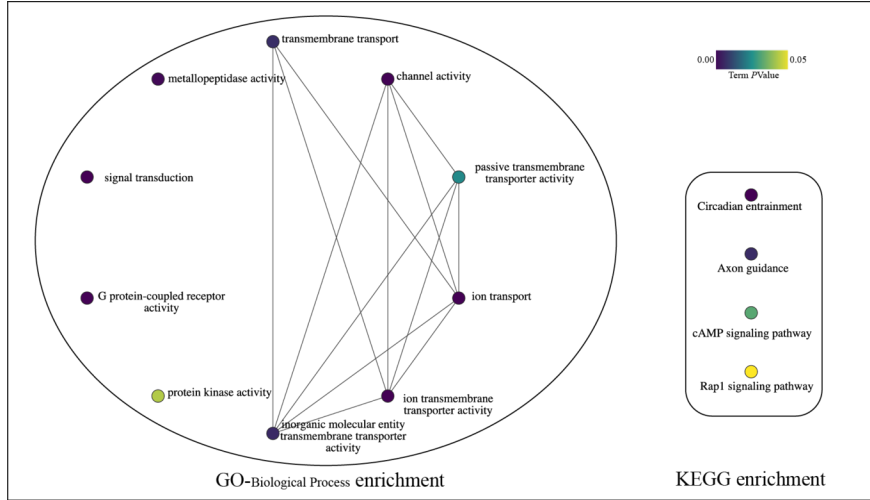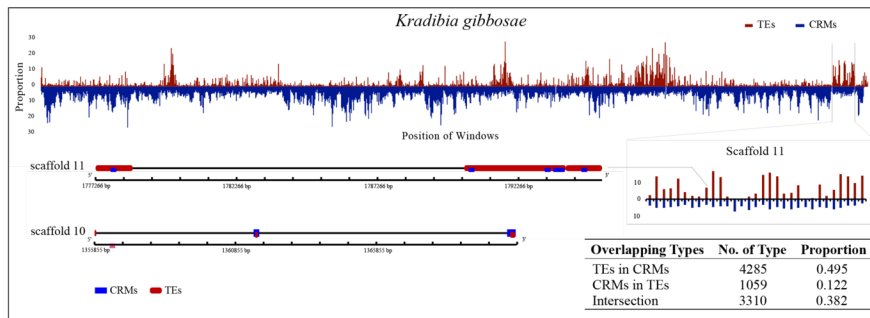
13

Figure 3: Population dynamics of fig wasps.

Figure 4:KEGG and GO network enrichment analysis of genes near classified TEs at the major peak of pollinators. The nodes were filtered using the parameter Term P- Value < 0.05.

Figure 5:Distribution and overlapping patterns between TEs and CRMs in Kgib. The histogram plotted the distribution of TEs and CRMs. Details of overlapping modes were shown in a partial sequence located at scaffold 10 and 11. Statistics of the overlapping mode were listed in the table.

Table 1 Details of fig wasps and TEs annotations.

Category	Species	Abbreviation	Host	Mating location	Oviposit
Pollinators	*Platyscapa corneri*	Pcor	*Ficus superba*	Inside	Inside
	Eupristina koningsbergeri	Ekon	*Ficus benjamina*	Inside	Inside
	Wiebesia pumilae	Wpum	*Ficus pumila*	Inside	Inside
	Dolichoris vasculosae	Dvas	*Ficus vasculosa*	Inside	Inside
	Kradibia gibbosae	Kgib	*Ficus gibbosa*	Inside	Inside
	Ceratosolen fusciceps	Cfus	*Ficus racemosa*	Inside	Inside
Non- pollinators	*Sycobia* sp.2	Sbsp	*Ficus benjamina*	Inside/ Outside	Outside
	Philotrypesis tridentata	Ptri	*Ficus benjamina*	Inside/ Outside	Outside
	Apocrypta bakeri	Abak	*C. solmsi* (*Ficus hispida*)	Inside	Outside
	Sycophaga agraensis	Sagr	*Ficus racemosa*	Inside	Outside
	Sycophila sp.2	Spsp	*Ficus benjamina*	Inside/ Outside	Outside

15

r : The nucleotide substitution rate per million year per site

μ : The neutral mutation rate per generation per site