Communications Biology

https://doi.org/10.1038/s42003-026-09699-4

Article in Press

Genomic landscape and genetic manipulation of an ectoparasitoid wasp, *Gregopimpla kuwanae*

Received: 9 April 2025

Accepted: 2 February 2026

Cite this article as: Gao, H., Li, Y., Chen, Y. *et al.* Genomic landscape and genetic manipulation of an ectoparasitoid wasp, *Gregopimpla kuwanae*. *Commun Biol* (2026). https://doi.org/10.1038/s42003-026-09699-4

Han Gao, Yijiangcheng Li, Yanli Chen, Xiaojing Liu, Mengying Fang, Shuyu Zhang, Jianhao Ding, Dalin Zhu, Anjiang Tan & Sheng Sheng

Genomic landscape and genetic manipulation of an ectoparasitoid wasp*, Gregopimpla kuwanae*

Han Gao[1,2], Yijiangcheng Li[1,2], Yanli Chen[1,2], Xiaojing Liu[1,2], Mengying Fang[1,2], Shuyu Zhang[1,2], Jianhao Ding[1,2], Dalin Zhu[1,2], Anjiang Tan[1,2,*], and Sheng Sheng[1,2,*]

[1] Jiangsu Key Laboratory of Sericultural and Animal Biotechnology, School of Biotechnology, Jiangsu University of Science and Technology, Zhenjiang 212100, China.

[2] Key Laboratory of Silkworm and Mulberry Genetic Improvement, Ministry of Agriculture and Rural Affairs, Sericultural Scientific Research Center, Chinese Academy of Agricultural Sciences, Zhenjiang 212100, China.

* To whom correspondence should be addressed: Prof. Anjiang Tan[*], and Sheng Sheng[*]

E-mail: atan@just.edu.cn (Anjiang Tan), parasitoids@163.com (Sheng Sheng)

Abstract: Parasitoid wasps are important biological control resources, yet their genetic manipulation has long been constrained by small body size and parasitization behavior, limiting their broader application in pest management. Here we report a chromosome-level genome assembly of the ectoparasitoid *Gregopimpla kuwanae* (322.87 Mb, 24 chromosomes), a relatively large species that parasitizes various lepidopteran pests. In the first part of this study, we established a foundational genomic resource and experimental platform by producing a high-quality genome and demonstrating the feasibility of functional genetics: RNA interference successfully silenced the *cinnabar* gene, while CRISPR/Cas9 editing generated *vestigial* knockout mutants, thus establishing *G. kuwanae* as a tractable system for gene manipulation. In the second part, we applied comparative genomics to identify lineage-specific gene-family expansions linked to parasitism, including venom-related genes, immune suppression factors, and detoxification enzymes (cytochrome P450s and UDP-glucosyltransferases), and we identified eight HGT candidates; one candidate (*JSFChr12G01362*) showed pre-feeding expression in females and caused increased adult mortality upon RNAi. Our study provides both the means and the candidates for mechanistic dissection of parasitoid adaptations, laying a foundation for the broader application of parasitoid wasps in sustainable biocontrol programs.

Key words: *Gregopimpla kuwanae*, Genome, RNA interference, CRISPR/Cas9

Introduction

The increasing global demand for food has intensified agricultural practices, often reducing natural enemy populations and triggering frequent pest outbreaks that cause substantial crop losses[1,2]. Although chemical pesticides remain widely used, their extensive application poses serious risks to ecosystems, food safety, and human health[3–5]. Biological control has therefore emerged as a sustainable alternative that exploits natural enemies to regulate pest populations. Among these, parasitoid wasps (Hymenoptera) are a highly diverse and ecologically important group of biocontrol agents[6]. Female parasitoid wasps lay eggs on or within host organisms, and their larvae consume host resources, ultimately leading to host death[7]. As key components of insect food webs, parasitoid wasps have long played an essential role in biological control programs[8].

Parasitoid wasps employ diverse oviposition strategies, with endoparasitoids developing inside hosts and ectoparasitoids developing externally[9,10] Although genome sequencing has been conducted for nearly 300 parasitoid species and behavioral studies have advanced[11–13], the genetic mechanisms underlying traits relevant to biological control remain poorly understood. Functional tools such as RNAi and CRISPR/Cas9 could address this gap, but their application is often hindered by small body size, endoparasitic development, and technical barriers to embryo manipulation[14–16]. Consequently, functional genetic studies are largely restricted to a few model species, including *Nasonia vitripennis* and *Habrobracon hebeto*[17,18]. *Gregopimpla kuwanae* is a relatively large ectoparasitoid that attacks lepidopteran pests[19–21], and its size and external development make it a promising system for establishing genetic protocols. However, genomic

resources for this species have been lacking.

Utilizing genomic data from closely related species to cluster and compare homologous genes has become a common approach to analyze biological traits of different species. For example, comparative genomics has revealed gene family expansions linked to adaptation in birds inhabiting the Tibetan Plateau[22] and in bats with specialized immune systems[23]. Furthermore, HGT has been increasingly recognized as a contributor to environmental adaptability and evolutionary innovation in insects[24]. Identifying expanded gene families and horizontally acquired genes in *G. kuwanae*, and exploring their potential functions, therefore represents an important objective.

Here, we present a chromosome-level genome assembly of *G. kuwanae* and establish efficient RNAi and CRISPR/Cas9 systems in this species. Using comparative genomic analyses, we identify rapidly evolving gene families and candidate HGT genes, and provide preliminary functional validation of one HGT-derived gene. Together, this work establishes *G. kuwanae* as both a genomic resource and an experimental model for elucidating the molecular basis of parasitism and advancing sustainable biological control.

Results

Chromosome-level Genome Assembly and Annotation of *G. kuwanae*

Sequencing on the DNBSEQ platform yielded a total of 44.34 Gb, with 43.75 Gb of clean reads (128X) after quality control filtering. K-mer survey analysis estimated the genome size at approximately 341,951,729 bp, with a heterozygosity rate of 0.82% and a repeat content of 38.24%, indicating a high-heterozygosity, low-repetition genome (Table 1 and Figure S1). PacBio HiFi

sequencing generated 6.7 Gb of clean data, providing a sequencing depth of approximately 20X. After assembly with Hifiasm (version 0.18.2) and redundancy removal, the preliminary assembled contig sequence reached a total length of 369.69 Mb. Following Hi-C sequencing, 143.02 Gb of clean data were obtained, and using Hi-C scaffolding, 322.87 Mb of sequences were mapped to 24 chromosomes, achieving a Hi-C mapping rate of 90.04%.

To assess the quality of the genome assembly, we divided the chromosome-level genome into bins of 500 kb. The number of Hi-C read pairs covering each pair of bins was used as a signal of interaction strength. The interaction intensity was higher for bin pairs located along the diagonal of the contact map (Figure 1A), indicating a good assembly quality. BUSCO analysis using the insecta_odb10 database showed that both the contig-level and chromosome-level assemblies contained 99.2% complete BUSCOs, reflecting high genome completeness.

To predict repetitive sequences, we employed Tandem Repeats Finder, RepeatMasker, and RepeatProteinMask software. Additionally, using the *de novo* method with LTRharvest and RepeatModeler, we identified a *de novo* TE library, which was further annotated with RepeatMasker. After redundancy removal, the total length of repetitive sequences was 161,865,851 bp, accounting for approximately 43.76% of the genome. Among the different repeat types, long terminal repeats (LTRs) retrotransposons represented the largest proportion, accounting for 34.31% of the genome (Figure 1B).

For gene structure prediction, we used EVM software to integrate *De novo* prediction, homologous annotation, and transcriptomic data. A total of 11,292 genes were annotated (Figure 1C and Figure

S2), with 99.71% of genes (11,259) supported by at least 80% of the corresponding evidence (Table S1). To evaluate the completeness of the gene set, we conducted a BUSCO analysis, which showed that the gene set fully covered 1,312 conserved proteins with a completeness of 95.98%. The predicted genes were annotated using multiple functional databases, and 86.36% of genes (9,752) were annotated in at least one database (Figure 1D).

RNAi in *G. kuwanae* with Prolonged Action and Significant Gene Suppression

The *cinnabar* (*cn*) gene was selected as the target for RNAi, with changes in the eye color of adult wasps used to assess the success of the RNAi. The injection was performed when the 6-day-old larvae completed feeding and their bodies turned gray. At this stage, the larvae had completed feeding and were detached from the host. Approximately three days after the injection, the larvae began spinning silk and pupating, with the pupal stage lasting around 10–12 days before the adult emerged from the cocoon. We counted the number of larvae before pupation, during pupation, adults, and individuals that exhibited red eye coloration (Table S2). No significant difference in survival rates was found between the ds*cn* and ds*GFP* injection groups. About two-thirds of the individuals successfully pupated into adults (Figure 2A). When assessing the RNAi efficiency, we found that the dsRNA had not yet exerted its effect at 2 days post-injection, but by 4 days, a significant decrease in *cn* gene expression was detected. The effect persisted, and at 10 days post-injection, a clear reduction in *cn* gene expression was still evident (Figure 2B). At this point, the eye pigment deposition in *G. kuwanae* was nearly complete. During the adult stage, we observed eye color changes, with approximately one-fourth of individuals in the ds*cn*-injected group

showing dark-red eyes, whereas the eyes of the ds*GFP*-injected group remained black (Figure 2C and Figure 2D).

Gene Knockout Experiment with High Hatching Rate in *G. kuwanae*

To assess the feasibility of gene knockout in *G. kuwanae*, the *vestigial* (*vg*) gene was selected as a marker. Successful knockout of the *vg* gene in *G. kuwanae* results in adults displaying wing malformations. Due to their small size and parasitoid lifestyle, CRISPR/Cas9-mediated gene knockout has been reported in only a few parasitoid wasp species. The eggs of *G. kuwanae* are 1.5 mm in length and are laid on the surface of the host, which greatly facilitates gene editing experiments on the eggs. We designed two sgRNAs targeting the first and second exons to disrupt the *vg* gene (Figure 3A). A total of 300 eggs were collected 1 hour after oviposition and injected within 2 hours. After 36 hours, over 80% (245) of the injected eggs hatched normally. However, due to competition for food and limited space during larval development, the common phenomenon in the gregarious parasitism of *G. kuwanae*, a substantial number of larvae failed to survive, and ultimately 52 pupae were obtained. From these pupae, 45 adults emerged, among which 5 individuals exhibited varying degrees of wing defects. Notably, these vg-mutant individuals included both males and females. Consistent with the sexual dimorphism of *G. kuwanae*, the smaller body size observed in the vg mutant M2 is attributable to its male sex, whereas mutants M1 and M3 are females and therefore larger in size. Sequencing of the edited regions revealed five distinct types of deletions (Figure 3B and Figure 3C). Apart from the changes in wing morphology, no other developmental or morphological abnormalities were observed in *G.*

kuwanae. These results demonstrate the feasibility of gene editing in *G. kuwanae* and highlight the excellent hatching rate, providing a solid foundation for further applications of gene editing.

Gene Family Expansion Analysis Reveals the Formation of Ectoparasitoid Behavior in *G. kuwanae*

Using 3,829 single-copy orthologous genes from 19 representative Hymenoptera species spanning all superfamilies[25,26], we reconstructed a phylogenetic tree (Figure 4). Divergence time estimation indicated that Hymenoptera originated approximately 250 million years ago, and that the Ichneumonidae family diverged around 120 million years ago, consistent with previous reports on Hymenoptera phylogeny and divergence times[25,26]. Gene family evolution inferred with CAFE revealed an overall trend of gene family contraction across Hymenoptera, with pronounced losses coinciding with the emergence of parasitoid and social lifestyles, suggesting adaptive streamlining. Compared with two endoparasitoid ichneumonids, *G. kuwanae* exhibited 327 expanded and 2,124 contracted gene families. Further screening based on the criterion of $P < 0.01$ identified 61 gene families in *G. kuwanae* that exhibited the most significant changes, referred to as rapidly evolving gene families. Among them, 18 gene families showed rapid expansion. Four were limited by database constraints and could not be accurately classified, two were annotated as transposon families, and the remaining 12 families had accurate annotations (Table 2).

These rapidly expanding gene families can be roughly classified into five categories based on their association with parasitization behavior: detoxify host immune molecules and chemicals (e.g., CYP450, UGT, carboxylesterases, etc.), regulate host immune responses (e.g., neprilysin, venom

metalloproteinase, hemolymph LPS-binding protein, etc.), degrade host tissues to promote wasp egg development (e.g., chymotrypsin, venom carboxylesterase, etc.), alter host gene expression to favor parasitoid wasp growth (e.g., insulin-like growth factor-binding protein, terminal uridylyltransferase, etc.), and optimize host energy resources to support pupal development (e.g., acyl-CoA delta-9 desaturase, etc.). While endoparasitoids rely on internal host resources, ectoparasitoids like *G. kuwanae* necessitate a rapid host-killing approach to enable successful egg attachment and development externally. These rapidly expanding gene families may serve as the molecular basis for the enhanced parasitoid efficiency observed in *G. kuwanae*.

Metabolic Detoxification Enzyme Family Contracted during Hymenoptera Evolution Occurred Small Expansion in *G. kuwanae*

We identified expanded metabolic detoxification enzyme families in G. kuwanae, including CYP450 and UGT. The G. kuwanae genome contains 85 CYP450 genes and 16 UGT genes. Regarding chromosomal arrangement, most members of these two metabolic detoxification gene families in *G. kuwanae* are arranged in tandem (Figure 5A). The CYP450 family contains 56 members, forming 15 tandem duplication clusters of varying sizes, while 11 UGT members are clustered in a large tandem duplication cluster. This is consistent with previous reports, as metabolic detoxification enzyme families in insects are often arranged in tandem duplications[27]. The formation or further expansion of these tandem duplication clusters has led to the expansion of the CYP450 and UGT families in *G. kuwanae*. Overall, the CYP450 family in Hymenoptera has shown a contraction trend during evolution, especially after the emergence of parasitization

behavior (Figure 5B). A similar contraction trend is also observed in the UGT family (Figure S3), which may be due to the reduction in environmental complexity faced by parasitoid wasps.

Subfamily classification based on *Drosophila* CYP450 annotations revealed that the expansion in *G. kuwanae* is mainly concentrated in the CYP2 and CYP3 subfamilies (Figure 5C and Figure S4). Transcriptomic analyses across nine developmental stages showed generally low expression of most detoxification genes, consistent with their inducible nature. Notably, two CYP4 members (*JSFChr14G02364* and *JSFChr21G05916*) were highly expressed across all stages (Figure 5D). Phylogenetic analysis identified them as homologs of *D. melanogaster Cyp4g1* and *Cyp4p*/*Cyp4ac*, respectively. *Cyp4g1* in Flybase is annotated as encoding the terminal oxidative decarbonylase in cuticular hydrocarbon biosynthesis within oenocytes, with expression during the entire developmental stage after 12 hours of egg deposition. *Cyp4p* and *Cyp4ac* are annotated as being potentially involved in insect hormone metabolism and the breakdown of synthetic insecticides, with expression observed in all stages except eggs[28]. The similar expression patterns suggest that *JSFChr14G02364* and *JSFChr21G05916* may participate in cuticular hydrocarbon biosynthesis and hormone metabolism like *D. melanogaster*.

Identification of HGT Genes and Preliminary Functional Validation

HGT is a common phenomenon in insects, occurring through processes such as parasitism, symbiosis, or viral infections[24]. Using the approach described by Shen et al[29], followed by manual curation, we identified a total of eight HGT genes in *G. kuwanae* (Table S3). Among them, six originated from bacteria, one from fungi, and one from plants. Two groups of HGT genes were of

particular interest. The first consists of a tandemly duplicated cluster (*JSFChr10G00365* - *JSFChr10G00368*), which is broadly distributed across Hymenoptera but exhibits a unique four-copy tandem duplication only in *G. kuwanae* (Figure S5). Given their wide presence among Hymenopteran species, these horizontally transferred genes may contribute to growth and developmental processes across the order, although the specific functional implications of their duplication in *G. kuwanae* remain to be determined.

The second gene of interest, *JSFChr12G01362*, showed markedly high expression in adult females prior to feeding (Figure 6A and 6B). RNAi targeting this gene was performed during the late pupal stage. Three days after injection, the wasps completed eclosion and initiated feeding. During the initial feeding and ovarian development period, the expression level of *JSFChr12G01362* decreased by over 70% (Figure 6C). We initially hypothesized that this reduction might impair the parasitoid learning behavior of *G. kuwanae*; however, behavioral assays revealed no significant difference in learning rate between the *dsJSFChr12G01362* and *dsGFP* groups. Notably, from day 6 onward, the mortality rate of the *dsJSFChr12G01362*-treated group increased by approximately 12%, reaching an ~18% difference from controls by day 12 (Figure 6D). This elevated mortality was also observed when wasps were fed with honey water instead of hosts. Aside from the increased mortality, no other significant differences were detected between treatment groups in terms of oviposition rate, parasitization success, or morphological structure of the midgut and venom gland. Collectively, these findings suggest that the horizontally transferred gene *JSFChr12G01362* plays an essential role in maintaining normal physiological activity in adult

female *G. kuwanae*, although its precise molecular function and regulatory mechanisms require further investigation.

Discussion

As valuable biological control agents, parasitoid wasps exhibit a wide range of species and diverse parasitoid strategies[10]. Understanding traits related to biological control is fundamental for the development of parasitoid wasps as biocontrol agents. However, the challenges associated with genetic manipulation of parasitoid wasps arise due to their small size and endoparasitoid behavior. Here, we present the chromosomal-level genome of *G. kuwanae*, a relatively large ectoparasitoid wasp. Its genome consists of 24 chromosomes, totaling 322.87 Mb. Previous reports suggest that the genome size of parasitoid wasps is significantly influenced by the content of transposable elements[13]. Different types of transposable elements play key roles in the evolutionary expansion of the genome size across various parasitoid wasp species. Previous studies indicate that genome size variation in parasitoid wasps is largely driven by transposable element content [30,31]. In *G. kuwanae*, LTR retrotransposons represent the largest proportion, accounting for 34.31% of the genome, and are the primary contributors to the genome size.

RNAi has been widely applied across various insect orders. It is particularly efficient and systemic in Coleopterans, especially in species from the families Tenebrionidae (e.g., *Tribolium castaneum*) and Chrysomelidae (e.g., *Leptinotarsa decemlineata* and *Diabrotica virgifera*)[16]. However, RNAi efficiency varies considerably in Lepidoptera, Diptera, Hymenoptera, and Hemiptera, particularly in Lepidoptera and Hemiptera[32–34]. Due to higher nuclease activity compared to other insects,

RNAi efficacy in these orders is often compromised[35]. In *G. kuwanae*, RNAi demonstrated stable and efficient gene silencing, with effects lasting over 10 days and reducing target gene expression by more than 80%. RNAi efficiency and stability in *G. kuwanae* are comparable to those of well-established insect models[36–39]. The CRISPR/Cas9 system has been successfully applied to gene knockout, knock-in, and expression regulation in various insect species[40,41]. However, due to their small body size and endoparasitoid lifestyle, most parasitoid wasps face mechanical damage during embryo injection. In our study, over 80% of *G. kuwanae* embryos successfully hatched 36 hours post-injection, indicating a high survival rate that greatly facilitates the generation of mutant lines. These features position *G. kuwanae* as a promising experimental model for functional studies in parasitoid wasps, complementing existing model systems such as *Nasonia vitripennis*[18].

Through comparison with various Hymenoptera species, we identified rapidly expanding gene families in *G. kuwanae*, including several genes potentially involved in regulating parasitoid behavior. Interestingly, although detoxification gene families such as CYP450 and UGT have shown an overall contraction trend during Hymenoptera evolution, we observed a modest expansion of these families in *G. kuwanae*. CYP450 and UGT families are classical metabolic detoxification enzyme families[42,43]. As an ectoparasitoid, *G. kuwanae* may encounter greater exposure to environmental xenobiotics and host-derived compounds than endoparasitoids, which develop internally within the host. This ecological difference may explain the retention and expansion of detoxification genes in *G. kuwanae*. In particular, members of the CYP2 subfamily, which play diverse roles and are widely regarded as environmental response genes in mammals[44],

have expanded in *G. kuwanae*. Such expansion may facilitate adaptation to its external parasitoid lifestyle. Notably, many detoxification genes in *G. kuwanae* are organized in tandem clusters, suggesting that tandem duplication has contributed to their recent expansion. Further functional studies will be needed to clarify whether these detoxification enzymes enhance the ecological adaptability and parasitoid efficiency of *G. kuwanae*. Additionally, venom metalloproteinases, common venom peptides found in hemotoxins[45], and venom carboxylesterase, which has been shown to possess lipolytic activity in *Bombus ignitus* and can cleave the glycerol backbone of phospholipids[46], were also expanded in *G. kuwanae*. These venom components likely facilitate rapid paralysis of the host and degradation of host tissues to promote larval growth. The insulin-like growth factor-binding protein (IGFBP) can competitively bind insulin-like peptides (ILPs) with the insulin receptor (InR)[47]. High expression of this protein in the venom of the ectoparasitoid *Scleroderma guani* suggests that it may inhibit host growth by suppressing the TOR and IIS/TOR signaling pathways, facilitating the growth of the parasitoid[48].

In addition to expanded gene families, we identified eight HGT genes in *G. kuwanae*, which is notably higher than the previously reported average of three HGT genes per hymenopteran species. Among these, six originated from bacteria, consistent with earlier findings that most insect-acquired HGTs are bacterial in origin[24]. Evolutionarily, the tandemly duplicated gene cluster *JSFChr10G00365 - JSFChr10G00368* is particularly intriguing. Homologs of this gene group can be traced back to the early parasitoid *Orussus abietinus* and are retained across most hymenopteran lineages, suggesting that they may have acquired essential biological functions within

Hymenoptera. Another bacterial-origin gene, *JSFChr12G01362*, displayed a distinct expression pattern, showing high expression in females before feeding, implying a role in post-emergence physiological regulation. RNAi-mediated knockdown of *JSFChr12G01362* did not affect parasitoid learning behavior or oviposition success but resulted in a significant increase in adult mortality, regardless of whether females were fed on hosts or honey water. These results suggest that *JSFChr12G01362* contributes to maintaining physiological homeostasis in adult females, supporting the idea that HGTs can be co-opted to enhance host adaptability and fitness.

In conclusion, this study establishes *G. kuwanae* as both a genomic resource and an experimental platform for parasitoid biology. By combining genome assembly, functional genetic tools, comparative analyses, and preliminary gene validation, we provide insights into the molecular basis of ectoparasitism and open avenues for advancing parasitoid wasps as sustainable biocontrol agents.

Methods

Insect Rearing

Laboratory colony of *G. kuwanae* was established from cocoons collected in mulberry fields at Jiangsu University of Science and Technology (Zhenjiang, China). *Galleria mellonella* larvae (Jiyuan Baiyun Biopesticide Co., Ltd., China) were used as hosts. Insects were maintained at 25 ± 2 °C, 60–80% relative humidity, and a 14 h light/10 h dark photoperiod. Adult wasps were kept in glass tubes and fed daily with 10% honey solution. For colony maintenance, mated females

parasitized cocooned prepupae of *G. mellonella*[20], and parasitized hosts were transferred to Petri dishes (8.2 cm × 1.4 cm) until parasitoid emergence.

Genomic Sequencing and Assembly

In this study, K-mer statistics were first used to survey the genome size. A short paired-end DNA library was constructed from a female adult *G. kuwanae* using the MGIEasy Universal DNA Library Prep Set (MGI-Shenzhen, China) and sequenced on the DNBSEQ platform (BGI-Shenzhen, China). Raw data were filtered using SOAPnuke-v2.1.04 software (parameters: -n 0.02 -l 20 -q 0.4 -i -G 2 --polyX 50 -Q 2 --seqType 0) to remove adaptor sequences and low-quality reads (Chen et al., 2018). K-mer frequency was rapidly counted using Jellyfish (v2.2.6) [50], and GenomeScope software was used to evaluate genome size, heterozygosity, repeat sequences, and sequencing depth based on K-mer data[51].

The genome was initially assembled using Pacific Biosciences (PacBio) high-fidelity (HiFi) sequencing. Libraries for PacBio HiFi long-read sequencing were prepared from genomic DNA and sequenced on the PacBio Sequel II platform using the CCS mode, with sequencing completed for one cell. Clean data were obtained after CCS (v4.0.0) filtering. Genome assembly was carried out using Hifiasm v0.18.2 (parameters: -o asm -t 32 --hom-cov 50)[52]. The preliminary assembly produced genome contig sequences, which were then refined by removing redundant sequences using Purge haplotigs4. Read-depth thresholds were determined from depth histograms (low=10, mid=40, high=120). To improve the assembly, Hi-C technology was used to obtain chromosome-level genomic information. The genome was digested with the MboI restriction enzyme, and a Hi-

C library was constructed, followed by sequencing on the DNBseq platform. Juicer was used to align paired-end sequencing data to the assembled genome, and the proportion of valid Hi-C contacts uniquely mapped to the genome was calculated for Hi-C library evaluation. Only these valid contacts could provide effective information for subsequent genome assembly. Hi-C contacts were then mounted using Juicer and 3D-DNA software with default parameter[53], and the assembly quality was evaluated using BUSCO v4 (parameters: -l insecta_odb10 -m genome)[54].

Genome Annotation

Genome repeats were annotated using a combination of homologous sequence alignment and de novo prediction methods. The homologous method relied on the RepBase library (https://www.girinst.org/repbase/) to identify sequences similar to known repeats using RepeatMasker and RepeatProteinMask[55], which were then classified. The de novo approach involved creating a repeat library with RepeatModeler and LTRharvest[56], followed by prediction with RepeatMasker. Additionally, Tandem Repeats Finder was used to identify tandem repeat sequences in the genome[57].

Gene structure annotation combined homologous prediction, de novo prediction, and transcriptome-assisted prediction. First, known protein-coding sequences from five homologous species (*Campoletis sonorensis*, *Diadromus collaris*, *N. vitripennis*, *Pteromalus puparum*, *Venturia canescens*), obtained from the NCBI (https://www.ncbi.nlm.nih.gov/datasets/genome/), were aligned to the *G. kuwanae* genome using GeMoMa to predict gene structures[58]. Augustus was then used to predict gene structures based on genomic features such as codon frequencies and

exon-intron distributions[59]. Additionally, transcriptome sequencing was performed on *G. kuwanae* at 10 developmental stages (egg, early larva, late larva, prepupa, female early pupa, female late pupa, male early pupa, male late pupa, female adult, and male adult) with three replicates per stage, totaling 30 samples, using the DNBSEQ platform (BGI-Shenzhen, China). These transcriptomic data were aligned to the assembled genome using HISAT to obtain transcript information. Finally, the predicted gene sets from all methods were integrated into a non-redundant and more complete gene set using EVM, and the quality of this gene set was assessed by comparing it to BUSCO[54]. To further obtain functional information, the annotated gene set was compared to the SwissProt, KEGG, KOG, and GO functional databases using Diamond[60], and InterProScan was used to query the InterPro database for functional annotations of all genes[61].

RNAi

Primers containing the *cn* gene sequences, along with the T7 polymerase promoter at the 5'-end of both forward and reverse primers, were used (Table S4). The PCR product was then used as a template for dsRNA synthesis using the MEGAscript RNAi Kit. Every 30 6-day-old larvae were used as a repeat for injection. 4.5 μL of dsRNA (*cn* as the positive control, with non-injections and injections of an equal volume of ds*GFP* as the negative control) was mixed with 0.5 μL of food coloring. RNAi injections were performed using the Narishige IM-400 microinjection system with the following parameters: injection pressure = 200 hPa; injection time = 0.3 s. RNAi efficiency was assessed by qRT-PCR using a pool of three individuals on the 2nd, 4th, 6th, 8th, and 10th days after dsRNA injection. The wasps were monitored daily, and both mortality and phenotypes were

recorded. At least three replications were conducted for each RNAi phenotype.

Gene Knockout

The coding sequence of *vg* was predicted and validated based on genome annotation results. Considering the PAM sequences, newly designed sgRNAs followed the 5'-CCNN$_{18}$CC-3' rule, and two 23-bp sgRNA targeting sites, named S1 and S2, were identified. The sgRNA target sequences were as follows: sgRNA-1 (S1): 5'-CCTTACCTCTATCAGAGGCCACC-3', sgRNA-2 (S2): 5'-CCTAGACCGACACCACCTCAACC-3'. sgRNA templates were transcribed using a T7 promoter and synthesized in vitro with the MAXIscript T7 Kit (Ambion, Austin, TX, USA) according to the manufacturer's instructions. Cas9 protein was purchased from Thermo Fisher Scientific (Catalog No. A36498). Fertilized eggs were collected within 1 hour of oviposition, and microinjection was carried out within 2 hours. Cas9 protein (200 ng/μL), along with sgRNA-1 (100 ng/μL) and sgRNA-2 (100 ng/μL), were co-injected into preblastodermal embryos. The injected eggs were then incubated in a humidified chamber at 25°C for 36 hours until hatching. The hatched larvae were subsequently reared on the surface of envenomed *G. mellonella* larvae that had just been attacked by *G. kuwanae* females. Morphological changes in wing development were observed in adult *G. kuwanae*, and Sanger sequencing of PCR-amplified genomic fragments spanning the two sgRNA target sites was used to detect CRISPR/Cas9-induced mutations in adults. PCR amplification was performed using primer pairs *vg*-TS1-F/*vg*-TS1-R and *vg*-TS2-F/*vg*-TS2-R, which are listed in Supplementary Table 4. Insertions and deletions (indels) were identified by aligning sequencing chromatograms to the wild-type sequence. Representative Sanger sequencing

chromatograms showing mutation patterns are Figure S6.

Evolutionary analysis

Nineteen sequenced hymenopteran species (Table S5) collected from InsectBase (https://www.insect-genome.com/genome) were used to infer gene orthology using OrthoFinder with default parameters[62]. First, TBTools was employed to remove alternative splicing events and retain the longest transcript for further analysis[63]. Gene orthology was then assessed using OrthoFinder to compare all protein sequences. Genes with a sequence similarity of 1^{e-5} or higher were grouped into the same gene family, classified as single-copy, multi-copy, or species-specific genes. Single-copy genes from each species were concatenated, and the concatenated alignment was subjected to multiple sequence alignment using MAFFT v7[64]. A phylogenetic tree was constructed using RAxML software[65], and divergence times between species were estimated with MCMCTree, calibrated using fossil data from the Timetree website (http://timetree.org/)[66,67].

Gene family expansion and contraction were analyzed using CAFE v4.2.1 (parameters: -p 0.01 -t 18)[68], based on a birth-death model to simulate gene family gains and losses during phylogenetic evolution. The transition rates of gene family size from parent to child nodes were calculated. An optimal birth-death parameter for the phylogenetic tree was estimated, considering divergence times of the 46 species. Gene families showing significant changes ($P < 0.01$) were classified as rapidly evolving. GO annotation and enrichment analysis of expanded gene families in *G. kuwanae* were conducted using Omicshare (https://www.omicshare.com/tools) to identify their biological functions. For the rapidly expanding gene families, specific classifications were performed by

aligning them with the UniProt database.

Identification of Expanded Detoxification Enzyme

Our study focused on the identification of expanded detoxification enzyme gene families (CYP450, UGT) in *G. kuwanae*. For the CYP450 and UGT families, known *Drosophila* sequences from Flybase were used as input. BLASTP searches (cut-off e-value = 1^{e-5}) were performed to identify potential candidates[69]. Domains were predicted using hmmsearch against the PFAM database. Candidate sequences containing characteristic domains were considered as members of the target families.

To further explore the evolutionary history of these gene families, phylogenetic tests were performed using the bootstrap method with 1000 replications to reconstruct maximum likelihood (ML) trees using IQ-TREE[70]. The best-fit tree model was determined by ModelFinder[71]. Evolutionary events, including gene gains and losses within these families, were analyzed using Notung[72]. MCScanX was used with default parameters to classify the duplication types of expanded genes within these families[73]. A whole-genome BLASTP analysis (max target seqs = 5, e-value cut-off = 1^{e-5}) was performed to determine homology with different genes in the *G. kuwanae* genome. Based on homology and chromosomal location, the genes were classified into various duplication types, including segmental and tandem duplications. All visualizations were performed using TBtools. Expression data from RNA-seq were also extracted for all family members, and heatmaps were generated using R to visualize temporal expression patterns.

Identification of HGT Genes

All annotated protein-coding sequences of the *G. kuwanae* genome were aligned against the NCBI non-redundant (NR) protein database using DIAMOND (cutoff e-value = 1^{e-10})[60]. Alien Index (AI) and Outgroup Percentage values were then calculated for each sequence using a script created by Shen et al[29]. A gene was considered a potential horizontally transferred gene when the top-scoring hit to non-metazoan sequences exceeded the top-scoring hit to metazoan sequences, and more than 80% of the aligned hits originated from non-metazoans. Candidate HGT genes were further validated by phylogenetic analyses. If the gene tree exhibited significant topological discordance with the species phylogeny, and the gene's evolutionary branch was firmly nested within the donor lineage, it was considered to have been acquired through horizontal gene transfer.

Statistics and Reproducibility

All statistical analyses were performed using GraphPad Prism version 8 (GraphPad Software, San Diego, CA, USA) unless otherwise stated. Sample sizes (n) for each experiment are indicated in the corresponding figure legends. For RNAi and gene knockout experiments, independent biological replicates were defined as groups of insects treated and analyzed independently on different occasions. RNAi experiments were repeated at least three times, and similar phenotypic and survival trends were observed across replicates. For qRT-PCR analyses, each data point represents the mean of three biologically independent samples, with each sample consisting of pooled individuals as described in the Methods. Survival data were analyzed using Kaplan–Meier survival analysis, and differences between survival curves were assessed using the Mantel-Cox (log-rank) test. For comparisons of gene expression levels between groups, independent-sample t-

tests were used. When multiple group comparisons were performed, appropriate multiple-comparisons correction was applied, and significance was indicated using different letters above bars; groups sharing the same letter were not significantly different ($P > 0.05$). Data are presented as mean ± SD, with individual data points shown where applicable.

Code availability

The software and parameters used in this study are described in the Methods section. No specific custom codes or scripts were utilized. Data processing was conducted according to the manuals and protocols provided with the respective software.

Data Availability statement

The genome assembly and all raw sequencing data generated in this study have been deposited in the NCBI under BioProject accession PRJNA1242975. This includes raw short-read sequencing data (SRX28190072), PacBio HiFi long-read data (SRX28190073), Hi-C data (SRX28190084 and SRX28190095), and RNA-seq data (SRX28190074-SRX28190083, SRX28190085-SRX28190094, SRX28190096-SRX28190111), as well as the final chromosome-level genome assembly of *G. kuwanae* (GCA_052576135.1). Source data underlying the figures of this manuscript can be found in Supplementary Data 1. All other data supporting the findings of this study are also available from the corresponding author upon reasonable request.

Author contributions

Hao Gao: Conceptualization, Methodology, Investigation, Writing-original draft; Yijiangcheng Li, Yanli Chen: Sample Collection, Experiment, Investigation; Mengying Fang, Shuyu Zhang, Hui Zhang: Sample Collection, Experiment; Jianhao Ding, Dalin Zhu: Investigation, Validation; Anjiang Tang, Sheng Sheng: Funding acquisition, Project administration, Supervision, Writing-review & editing.

Competing Interests

The authors declare no competing interests.

Acknowledgements

This work was supported by the National Science Foundation of China (31925007 and 32400380), the Natural Science Foundation of Jiangsu Province (BK20241861).

Reference

1. Beddington, J. Food security: contributions from science to a new and greener revolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 61-71 (2010).
2. Oerke, E.-C. Crop losses to pests. *The Journal of Agricultural Science* **144**, 31–43 (2006).
3. Barzman, M. *et al.* Eight principles of integrated pest management. *Agron. Sustain. Dev.* **35**, 1199–1215 (2015).
4. Damalas, C. A. & Eleftherohorinos, I. G. Pesticide Exposure, Safety Issues, and Risk Assessment Indicators. *International Journal of Environmental Research and Public Health* **8**, 1402–1419 (2011).
5. Mancini, F., Woodcock, B. A. & Isaac, N. J. B. Agrochemicals in the wild: Identifying links between pesticide use and declines of nontarget organisms. *Current Opinion in Environmental Science & Health* **11**, 53–58 (2019).
6. Blaimer, B. B. *et al.* Key innovations and the diversification of Hymenoptera. *Nat Commun* **14**, 1212 (2023).
7. Burke, G. R. & Sharanowski, B. J. Parasitoid wasps. *Current Biology* **34**, R483–R488 (2024).
8. Fei, M., Gols, R. & Harvey, J. A. The Biology and Ecology of Parasitoid Wasps of Predatory Arthropods. *Annual Review of Entomology* **68**, 109–128 (2023).
9. Chen, X. & van Achterberg, C. Systematics, Phylogeny, and Evolution of Braconid Wasps: 30 Years of Progress. *Annual Review of Entomology* **64**, 335–358 (2019).

10. Polaszek, A. & Vilhemsen, L. Biodiversity of hymenopteran parasitoids. *Current Opinion in Insect Science* **56**, 101026 (2023).
11. Strand, M. R. Teratocytes and their functions in parasitoids. *Current Opinion in Insect Science* **6**, 68–73 (2014).
12. Wang, Y. *et al.* Symbiotic bracovirus of a parasite manipulates host lipid metabolism via tachykinin signaling. *PLOS Pathogens* **17**, e1009365 (2021).
13. Ye, X., Yang, Y., Zhao, X., Fang, Q. & Ye, G. The state of parasitoid wasp genomics. *Trends in Parasitology* **40**, 914–929 (2024).
14. Hammond, A. *et al.* A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. *Nature Biotechnoogyl* **34**, 78–83 (2016).
15. Xu, H. *et al.* Comparative Genomics Sheds Light on the Convergent Evolution of Miniaturized Wasps. *Molecular Biology and Evolution* **38**, 5539–5554 (2021).
16. Zhu, K. Y. & Palli, S. R. Mechanisms, Applications, and Challenges of Insect RNA Interference. *Annual Review of Entomology* **65**, 293–311 (2020).
17. Bai, X. *et al.* CRISPR/Cas9-mediated mutagenesis of the white gene in an ectoparasitic wasp, *Habrobracon hebetor*. *Pest Management Science* **80**, 1219–1227 (2024).
18. Li, M. *et al.* Generation of heritable germline mutations in the jewel wasp *Nasonia vitripennis* using CRISPR/Cas9. *Scientific Reports* **7**, 901 (2017).
19. Li, Y. *et al.* Identification of candidate chemosensory genes by antennal transcriptome analysis in an ectoparasitoid wasp. *Journal of Applied Entomology* **146**, 335–351 (2022).
20. Ueno, T. Oviposition and Development in *Gregopimpla kuwanae* Viereck (Hymenoptera: Ichneumonidae), a Gregarious Ectoparasitoid Wasp Attacking the Rice Skipper *Parnara guttata*. *Journal of Insects* **2016**, e4706376 (2016).
21. Zhou, H., Yu, Y., Tan, X., Chen, A. & Feng, J. Biological control of insect pests in apple orchards in China. *Biological Control* **68**, 47–56 (2014).
22. Qu, Y. *et al.* Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan plateau. *Nature Communications* **4**, 2071 (2013).
23. Pavlovich, S. S. *et al.* The Egyptian Rousette Genome Reveals Unexpected Features of Bat Antiviral Immunity. *Cell* **173**, 1098-1110.e18 (2018).
24. Li, Y. *et al.* HGT is widespread in insects and contributes to male courtship in lepidopterans. *Cell* **185**, 2975-2987.e10 (2022).
25. Branstetter, M. G. *et al.* Phylogenomic Insights into the Evolution of Stinging Wasps and the Origins of Ants and Bees. *Current Biology* **27**, 1019–1025 (2017).
26. Peters, R. S. *et al.* Evolutionary History of the Hymenoptera. *Current Biology* **27**, 1013–1018 (2017).
27. Shi, Y. *et al.* Divergent amplifications of CYP9A cytochrome P450 genes provide two noctuid pests with differential protection against xenobiotics. *Proceedings of the National Academy of Sciences* **120**, e2308685120 (2023).

28. Thurmond, J. *et al.* FlyBase 2.0: the next generation. *Nucleic Acids Research* **47**, D759–D765 (2019).
29. Shen, X.-X. *et al.* Tempo and Mode of Genome Evolution in the Budding Yeast Subphylum. *Cell* **175**, 1533-1545.e20 (2018).
30. Ye, X. *et al.* Genomic signatures associated with maintenance of genome stability and venom turnover in two parasitoid wasps. *Nature Communication* **13**, 6417 (2022).
31. Yang, Y. *et al.* Genome of the pincer wasp *Gonatopus flavifemur* reveals unique venom evolution and a dual adaptation to parasitism and predation. *BMC Biology* **19**, 145 (2021).
32. Christiaens, O., Swevers, L. & Smagghe, G. DsRNA degradation in the pea aphid (*Acyrthosiphon pisum*) associated with lack of response in RNAi feeding and injection assay. *Peptides* **53**, 307–314 (2014).
33. Guo, X., Wang, Y., Sinakevitch, I., Lei, H. & Smith, B. H. Comparison of RNAi knockdown effect of tyramine receptor 1 induced by dsRNA and siRNA in brains of the honey bee, *Apis mellifera*. *Journal of Insect Physiology* **111**, 47–52 (2018).
34. Singh, I. K., Singh, S., Mogilicherla, K., Shukla, J. N. & Palli, S. R. Comparative analysis of double-stranded RNA degradation and processing in insects. *Scientific Reports* **7**, 17059 (2017).
35. Wang, K. *et al.* Variation in RNAi efficacy among insect species is attributable to dsRNA degradation *in vivo*. *Insect Biochemistry and Molecular Biology* **77**, 1–9 (2016).
36. Laudani, F. *et al.* RNAi-mediated gene silencing in *Rhynchophorus ferrugineus* (Oliver) (Coleoptera: Curculionidae). *Open Life Sciences* **12**, 214–222 (2017).
37. Prentice, K. *et al.* RNAi-based gene silencing through dsRNA injection or ingestion against the African sweet potato weevil *Cylas puncticollis* (Coleoptera: Brentidae). *Pest Management Science* **73**, 44–52 (2017).
38. Rangasamy, M. & Siegfried, B. D. Validation of RNA interference in western corn rootworm *Diabrotica virgifera* virgifera LeConte (Coleoptera: Chrysomelidae) adults. *Pest Management Science* **68**, 587–591 (2012).
39. Rodrigues, T. B., Dhandapani, R. K., Duan, J. J. & Palli, S. R. RNA interference in the Asian Longhorned Beetle: Identification of Key RNAi Genes and Reference Genes for RT-qPCR. *Scientific Reports* **7**, 8913 (2017).
40. Ghosh, S., Tibbit, C. & Liu, J.-L. Effective knockdown of *Drosophila* long non-coding RNAs by CRISPR interference. *Nucleic Acids Research* **44**, e84 (2016).
41. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588 (2015).
42. Nauen, R., Bass, C., Feyereisen, R. & Vontas, J. The Role of Cytochrome P450s in Insect Toxicology and Resistance. *Annual Review of Entomology* **67**, 105–124 (2022).
43. Zhu, S. *et al.* Expression of a viral ecdysteroid UDP-glucosyltransferase enhanced the insecticidal activity of the insect pathogenic fungus. *Pest Management Science* **80**, 4915–4923 (2024).

44. Feyereisen, R. 8 - Insect CYP Genes and P450 Enzymes. in *Insect Molecular Biology and Biochemistry* (ed. Gilbert, L. I.) 236–316 (Academic Press, San Diego, 2012).
45. Honutagi, R. M. *et al.* Protein-protein interaction of LDH and CRP-1 with hematotoxin snake venom proteins of all species of snake: An in silico approach. *International Journal of Health Sciences* **17**, 10 (2023).
46. Deng, Y. *et al.* Lipolytic Activity of a Carboxylesterase from Bumblebee (*Bombus ignitus*) Venom. *Toxins* **13**, 239 (2021).
47. Andersen, A. S., Hansen, P. H., Schäffer, L. & Kristensen, C. A New Secreted Insect Protein Belonging to the Immunoglobulin Superfamily Binds Insulin and Related Peptides and Inhibits Their Activities *. *Journal of Biological Chemistry* **275**, 16948–16953 (2000).
48. Zhu, J.-Y. Deciphering the main venom components of the ectoparasitic ant-like bethylid wasp, *Scleroderma guani*. *Toxicon* **113**, 32–40 (2016).
49. Chen, Y. *et al.* SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *GigaScience* **7**, gix120 (2018).
50. Hoover, A. & Miller, L. A numerical study of the benefits of driving jellyfish bells at their natural frequency. *Journal of Theoretical Biology* **374**, 13–25 (2015).
51. Vurture, G. W. *et al.* GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics* **33**, 2202–2204 (2017).
52. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nature Methods* **18**, 170–175 (2021).
53. Durand, N. C. *et al.* Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Systems* **3**, 95–98 (2016).
54. Seppey, M., Manni, M. & Zdobnov, E. M. BUSCO: Assessing Genome Assembly and Annotation Completeness. in *Gene Prediction: Methods and Protocols* (ed. Kollmar, M.) 227–245 (Springer, New York, NY, 2019).
55. Tarailo-Graovac, M. & Chen, N. Using RepeatMasker to Identify Repetitive Elements in Genomic Sequences. *Current Protocols in Bioinformatics* **25**, 4.10.1-4.10.14 (2009).
56. Ellinghaus, D., Kurtz, S. & Willhoeft, U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics* **9**, 18 (2008).
57. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* **27**, 573–580 (1999).
58. Keilwagen, J., Hartung, F. & Grau, J. GeMoMa: Homology-Based Gene Prediction Utilizing Intron Position Conservation and RNA-seq Data. in *Gene Prediction: Methods and Protocols* (ed. Kollmar, M.) 161–177 (Springer, New York, NY, 2019).
59. Stanke, M. *et al.* AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Research* **34**, W435–W439 (2006).

60. Hernández-Salmerón, J. E. & Moreno-Hagelsieb, G. Progress in quickly finding orthologs as reciprocal best hits: comparing blast, last, diamond and MMseqs2. *BMC Genomics* **21**, 741 (2020).
61. Quevillon, E. *et al.* InterProScan: protein domains identifier. *Nucleic Acids Research* **33**, W116-120 (2005).
62. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology* **20**, 238 (2019).
63. Chen, C. *et al.* TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Molecular Plant* **13**, 1194–1202 (2020).
64. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**, 3059–3066 (2002).
65. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
66. dos Reis, M. Dating Microbial Evolution with MCMCtree. in *Environmental Microbial Evolution: Methods and Protocols* (ed. Luo, H.) 3–22 (Springer US, New York, NY, 2022).
67. Kumar, S. *et al.* TimeTree 5: An Expanded Resource for Species Divergence Times. *Molecular Biology and Evolution* **39**, msac174 (2022).
68. Abramova, A., Osińska, A., Kunche, H., Burman, E. & Bengtsson-Palme, J. CAFE: a software suite for analysis of paired-sample transposon insertion sequencing data. *Bioinformatics* **37**, 121–122 (2021).
69. Johnson, M. *et al.* NCBI BLAST: a better web interface. *Nucleic Acids Research* **36**, W5-9 (2008).
70. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* **32**, 268–274 (2015).
71. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **14**, 587–589 (2017).
72. Chen, K., Durand, D. & Farach-Colton, M. NOTUNG: A Program for Dating Gene Duplications and Optimizing Gene Family Trees. *Journal of computational biology : a journal of computational molecular cell biology* **7**, 429-447 (2004)
73. Wang, Y. *et al.* MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research* **40**, e49 (2012).

Tables

Table 1. Statistics of the Repeat Sequence in *G. kuwanae*

Type		Length(bp)	% of genome
Retro	LTR/Copia	1,209,311	0.32694
	LTR/Gypsy	9,925,877	2.68352
	LTR/Other	115,787,663	31.30386
	SINE	149,265	0.04035
	LINE	4,296,764	1.16165
	Other	0	0
DNA	EnSpm	2,362,187	0.63863
	Harbinger	91,818	0.02482
	hAT	1,900,248	0.51374
	Helitron	2,741,806	0.74126
	Mariner	2,630,808	0.71125
	MuDR	1,120,239	0.30286
	P	121,295	0.03279
	Other	19,188,374	5.18769
Other	-	1,010,446	0.27318
Unknown	-	31,464,627	8.50664
Total	-	161,865,851	43.76137

Table 2. Overview of Rapidly Evolving Gene Families in *G. kuwanae*

Orthogroup	Expanded Number	Gene Number	Annotation
OG0000016	+5*	13	UDP-glucosyltransferase (UGT)
OG0000067	+5*	10	Chymotrypsin
OG0000100	+5*	13	Perhaps Transposable element P transposase
OG0000134	+5*	7	Perhaps Transposable element P transposase
OG0000154	+4*	8	Cytochrome P450 (CYP450)
OG0000159	+5*	10	Neprilysin
OG0000200	+4*	8	Carboxylesterase family
OG0000208	+3*	5	NA
OG0000228	+4*	8	Venom metalloproteinase
OG0000269	+4*	6	NA
OG0000275	+3*	5	NA
OG0000334	+5*	8	Terminal uridylyltransferase
OG0000396	+16*	18	Protein roadkill
OG0000397	+10*	13	Venom carboxylesterase
OG0000561	+13*	15	Insulin-like growth factor-binding protein

OG0001315	+14*	17	Hemolymph lipopolysaccharide-binding protein
OG0008038	+4*	6	NA
OG0010218	+8*	9	Acyl-CoA Delta-9 desaturase

Figure Legends

Figure 1. Overview of *G. kuwanae* Genome Assembly and Annotation. (A) Hi-C contact heatmap for 24 chromosomes. (B) Relationship between the variation degree of different types of repetitive sequences and their proportion in the genome. (C) Circos plot of the genome: each segment in the outermost circle represents a chromosome; within each chromosome, every 1 MB sequence is treated as a unit. Panels a-d show: (a) scatter plot of *G. kuwanae* repetitive sequence distribution; (b) GC content across different regions of *G. kuwanae*; (c-d) heatmap and histogram of gene locations in *G. kuwanae*. (D) Annotation status of all genes in various databases.

Figure 2. RNAi Results for *cn*. (A) Kaplan-Meier survival curves of individuals injected with ds*cn*, ds*GFP*, or WT (n = 30). Survival differences were analyzed using the Mantel-Cox (log-rank) test. (B) Expression levels of *cn* gene on different days post-injection. Statistical significance was assessed using independent-sample t-tests; data are shown as mean ± SD with individual data points (n = 3). *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. (C-D) Changes in eye color after ds*cn* injection; white triangles indicate changes in eye color post-injection.

Figure 3. CRISPR/Cas9-Mediated *vg* Gene Knockout. (A) Designed sgRNA target sites. (B) Mutation status after knockout. (C) Wing morphology changes post-knockout.

Figure 4. Evolution of Gene Families in *G. kuwanae*. (A) Expansion and contraction of gene families. The evolutionary tree on the right is constructed using single-copy orthologous genes from 19 selected species; numbers on nodes and branches represent the number of gene families that have expanded or contracted. The bar chart on the right indicates the number of genes in each category. (B) GO annotation bar chart of expanded gene families in *G. kuwanae*. (C) KEGG pathway enrichment bubble chart of expanded gene families in *G. kuwanae*.

Figure 5. Information on Rapidly Expanding CYP450 and UGT Families in *G. kuwanae*. (A) Chromosomal distribution of CYP450 and UGT families; blue represents UGT family, black represents P450 family. Genes with tandem duplications are highlighted with red boxes. (B) Phylogenetic classification of CYP subfamilies. (C) Evolution of CYP450 family in Hymenoptera; numbers at nodes represent gene birth/death events. (D) Expression heatmap of each member of CYP450 and UGT families in different stages of *G. kuwanae*.

Figure 6. Identification, expression, and RNAi analysis of the HGT gene *JSFChr12G01362*. (A) Phylogenetic tree constructed from the top 250 BLAST hits of *JSFChr12G01362* against the NCBI non-redundant protein database. Green branches represent bacterial sequences, while red branches

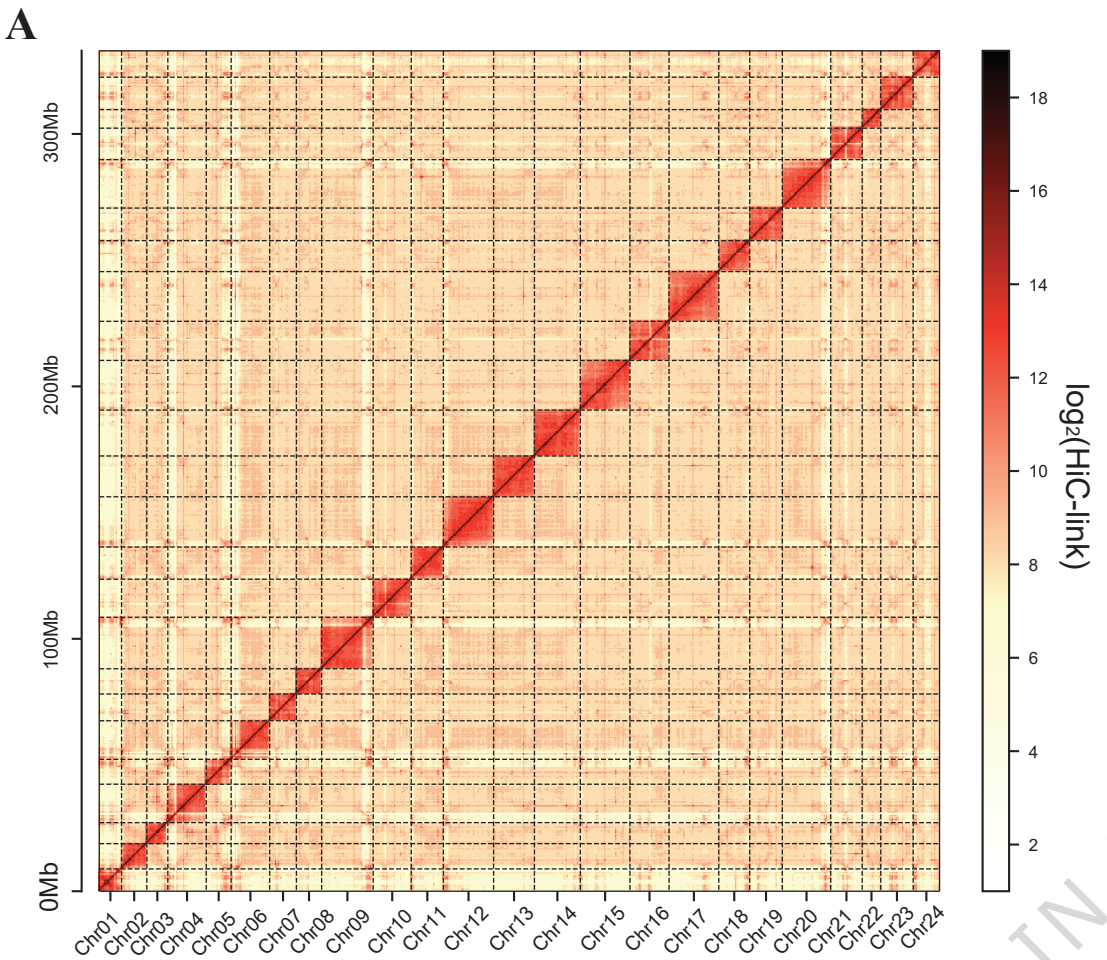
represent insect-derived sequences. (B) Developmental and tissue-specific expression profiles of *JSFChr12G01362*. Different letters above bars indicate statistically significant differences among groups ($P < 0.05$, independent-sample *t*-tests with multiple-comparisons correction), whereas identical letters indicate no significant difference ($P > 0.05$). Data are shown as mean ± SD, with individual data points overlaid; n = 3 biological replicates per group. (C) Relative expression levels of *JSFChr12G01362* at different time points after injection of ds*JSFChr12G01362* and ds*GFP*. Statistical significance was assessed using independent-sample t-tests; data are shown as mean ± SD with individual data points (n = 3). (D) Kaplan-Meier survival curves of adult females following ds*JSFChr12G01362* and ds*GFP* injection under two feeding conditions - host feeding and honey water feeding (n = 50). Survival differences were analyzed using the Mantel-Cox (log-rank) test. Asterisks indicate significance levels: ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$.

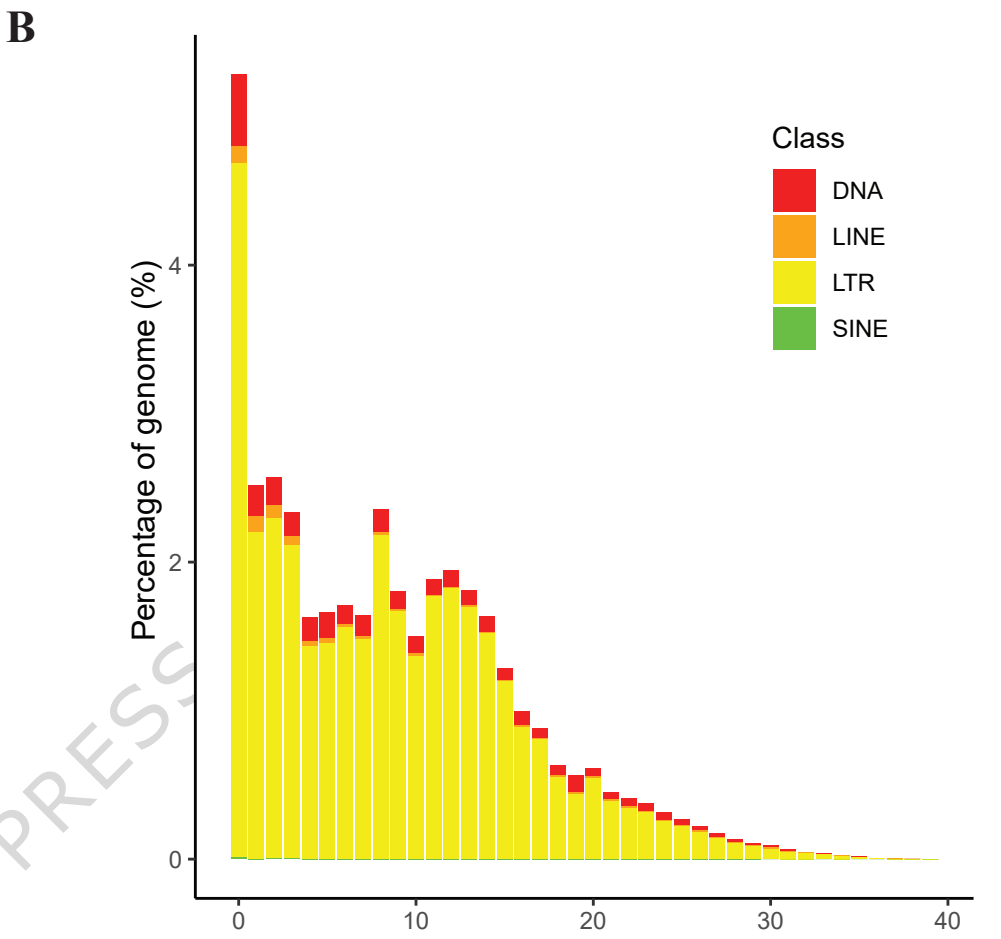
Editorial summary

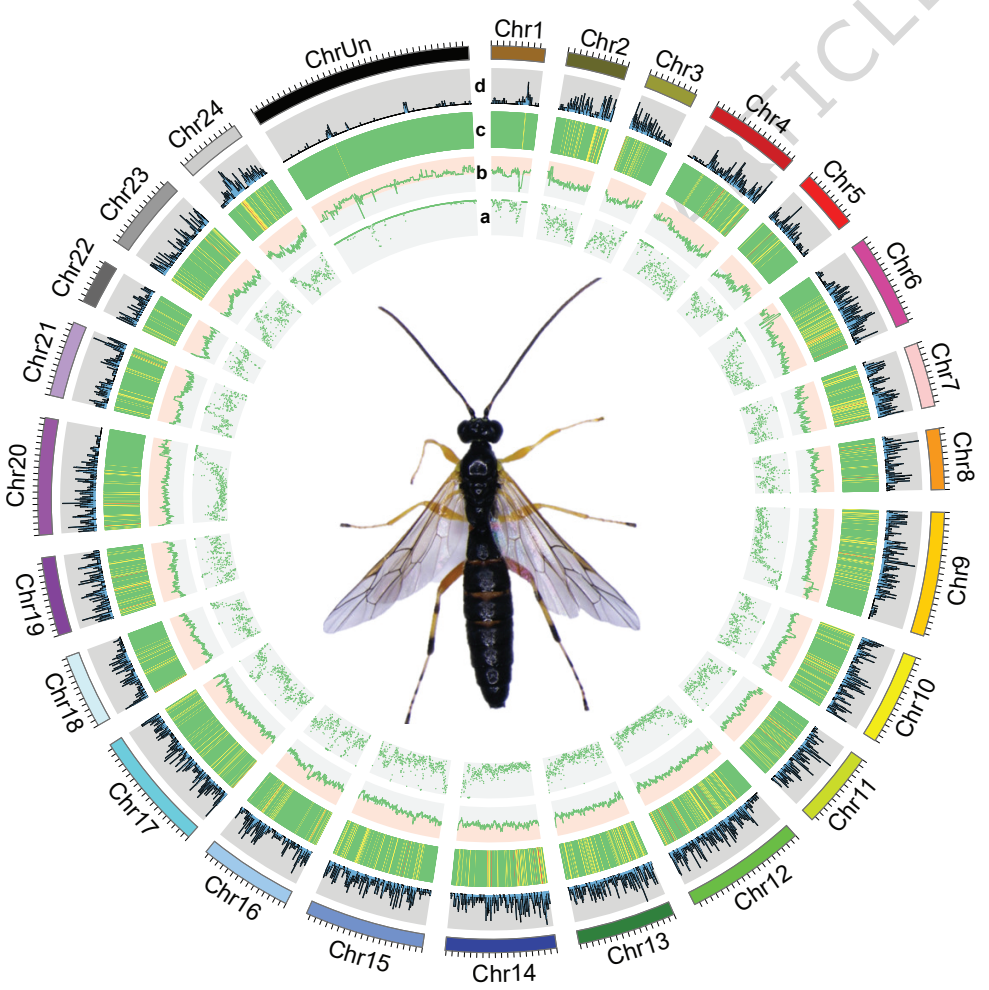
A high-quality genome assembly and RNAi/CRISPR toolkit establish Gregopimpla kuwanae as a tractable parasitoid model and uncover genomic features linked to ectoparasitic adaptation.

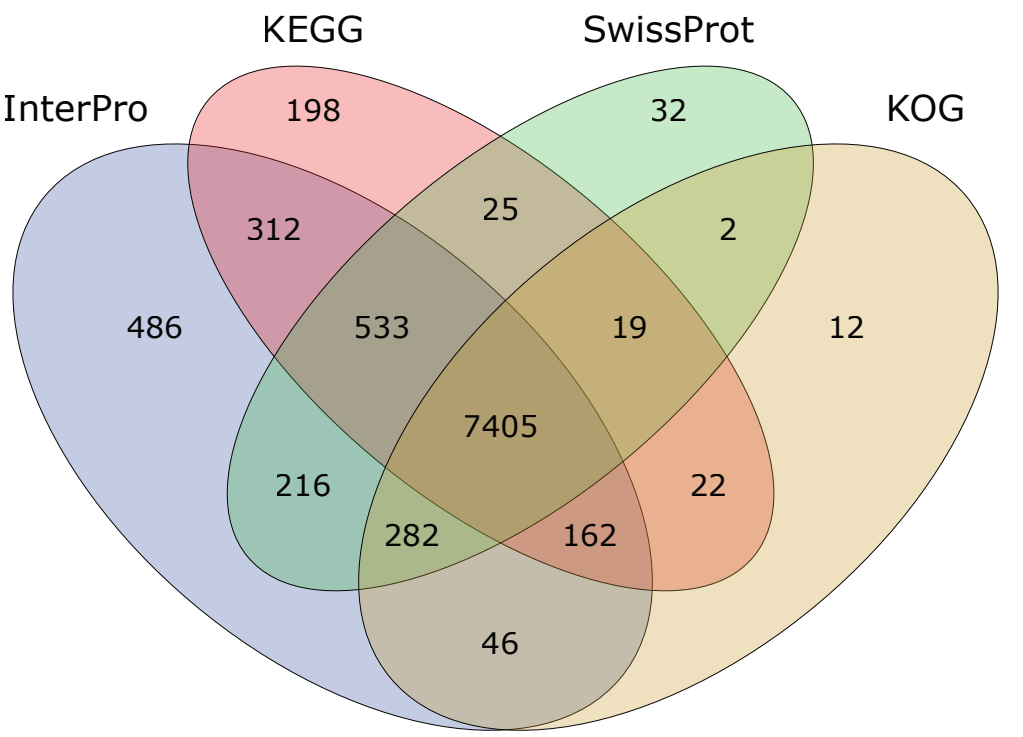
Peer review information

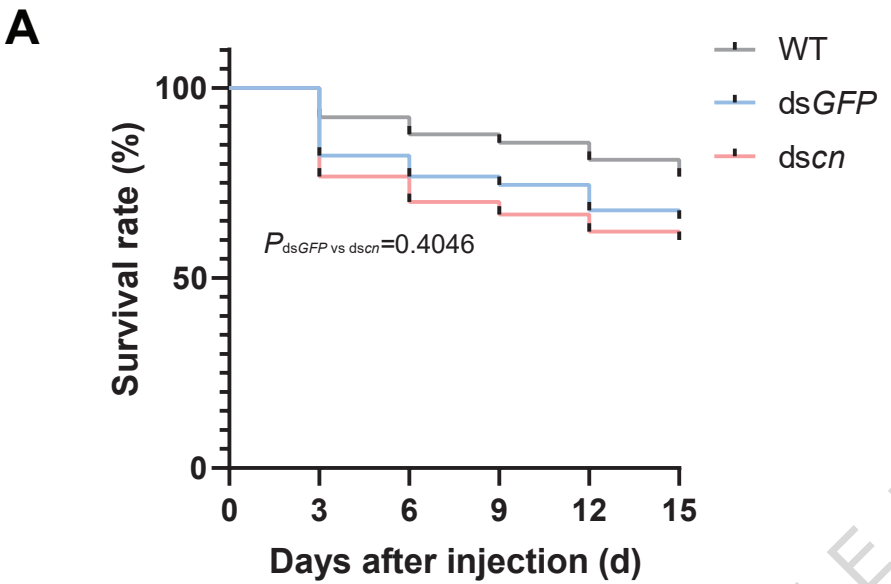
Communications Biology thanks Xinhai Ye, Yaohui Wang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Madhava Meegaskumbura & Rosie Bunton-Stasyshyn. A peer review file is available.

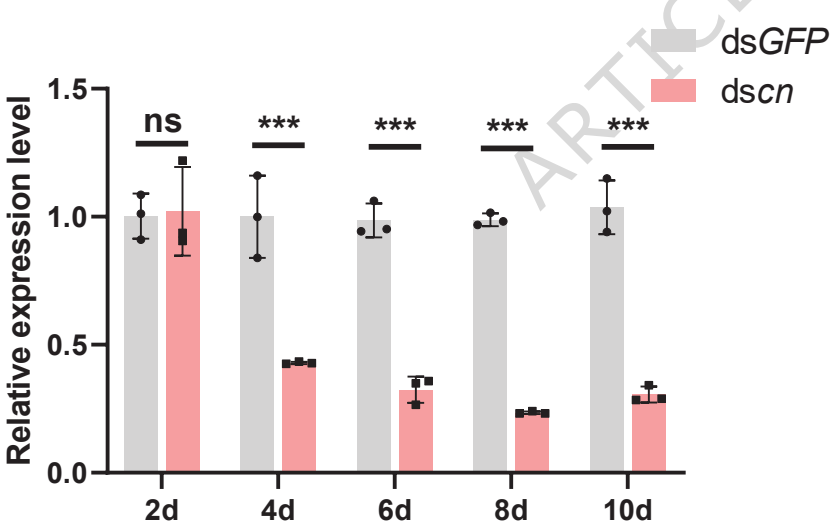


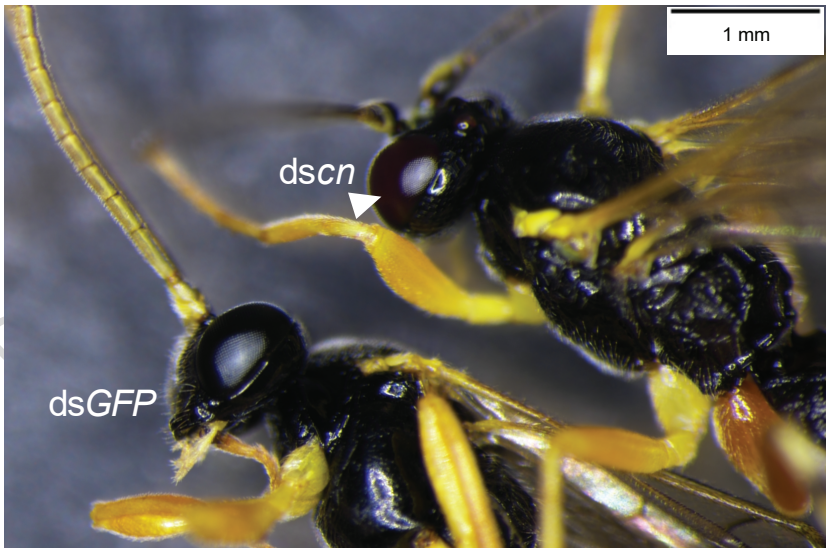


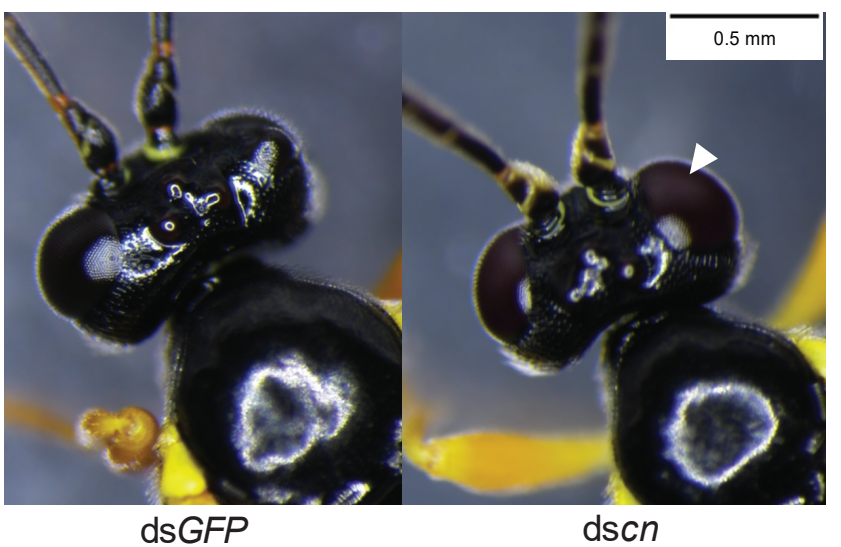




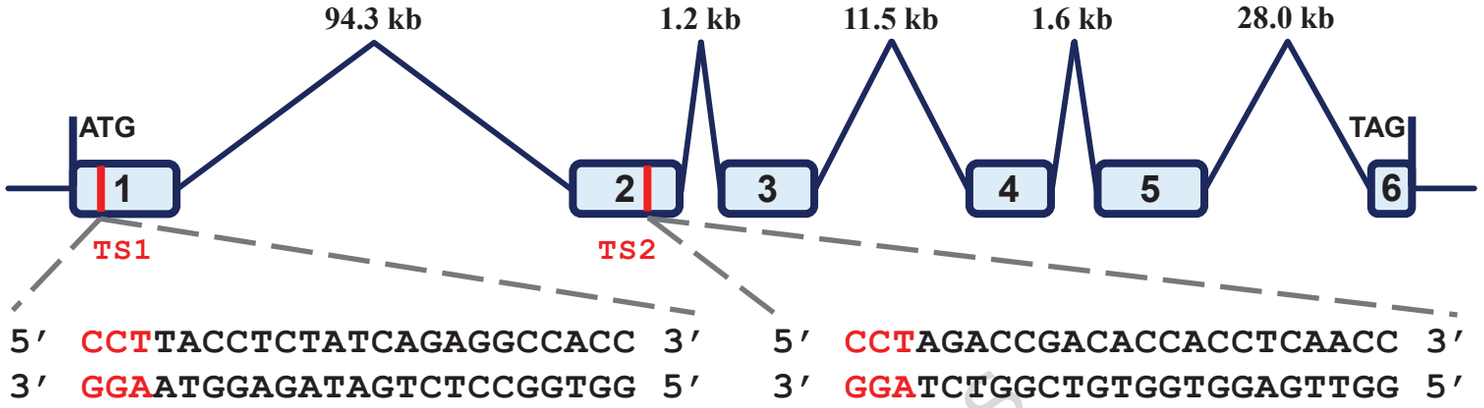








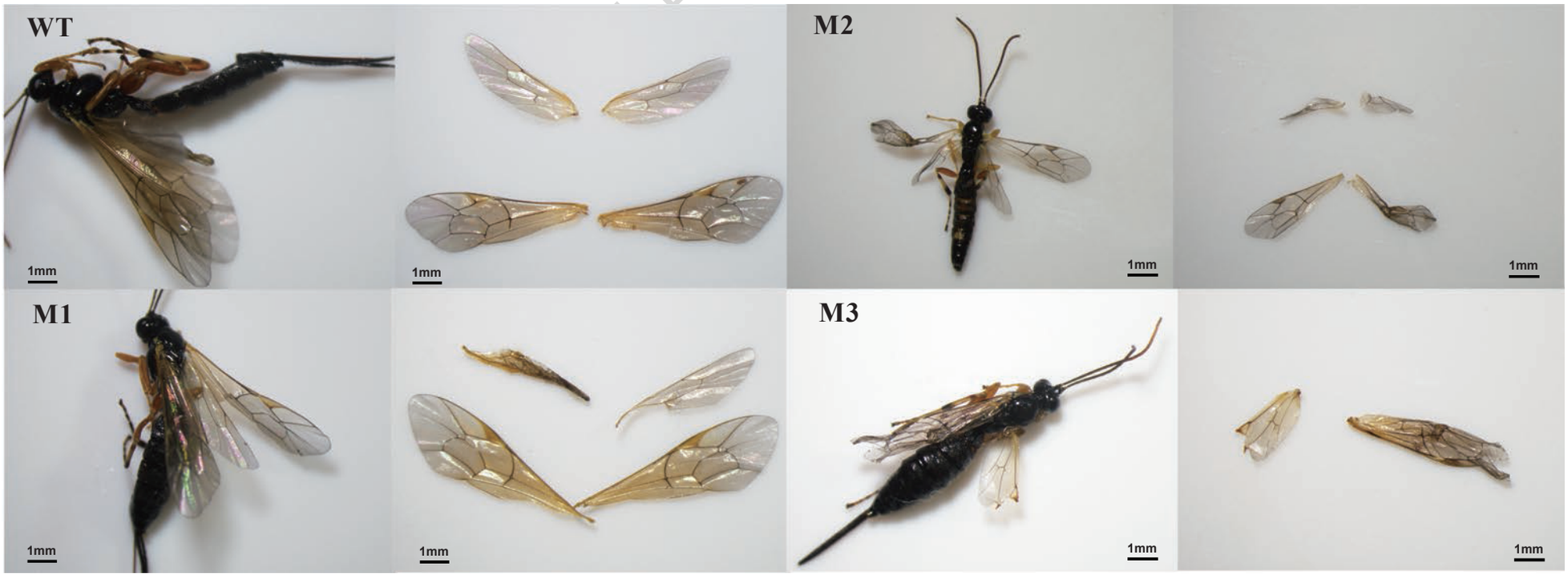
A

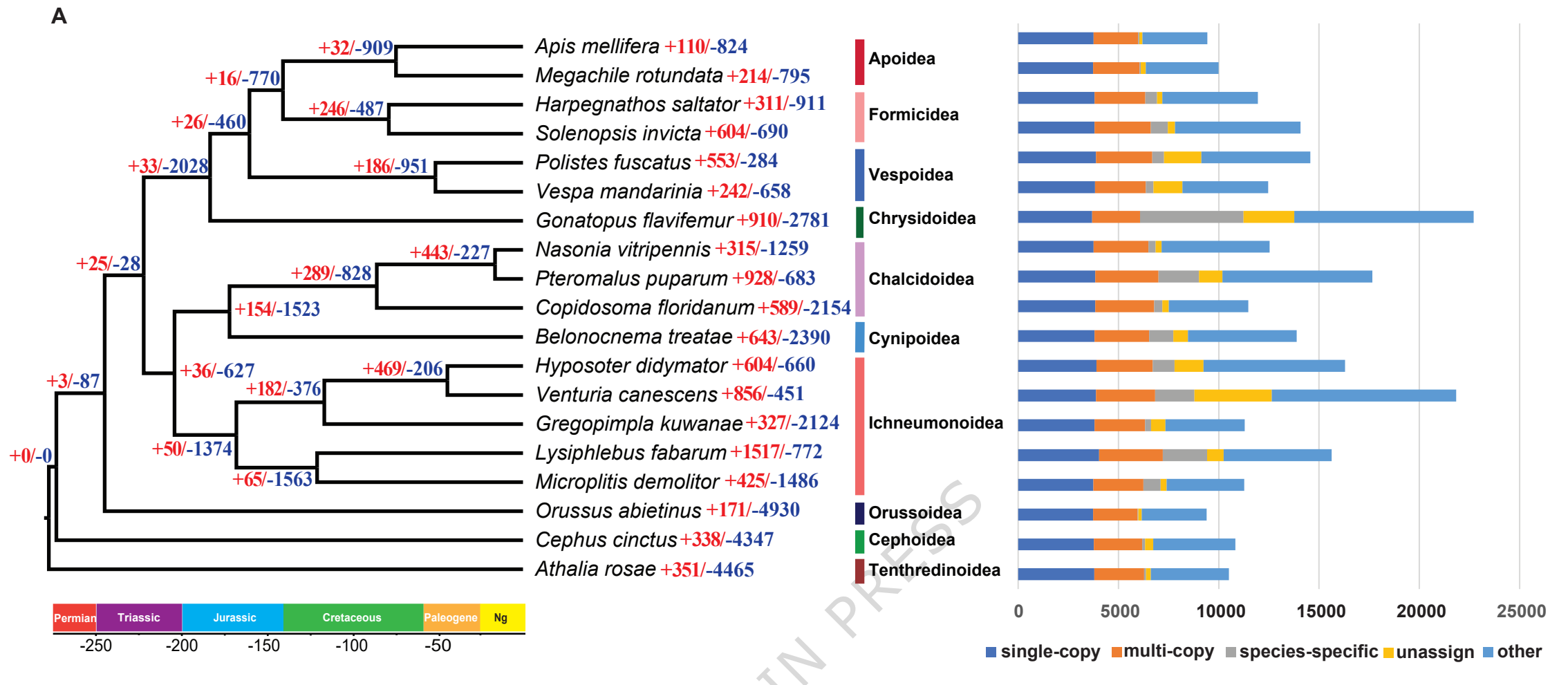


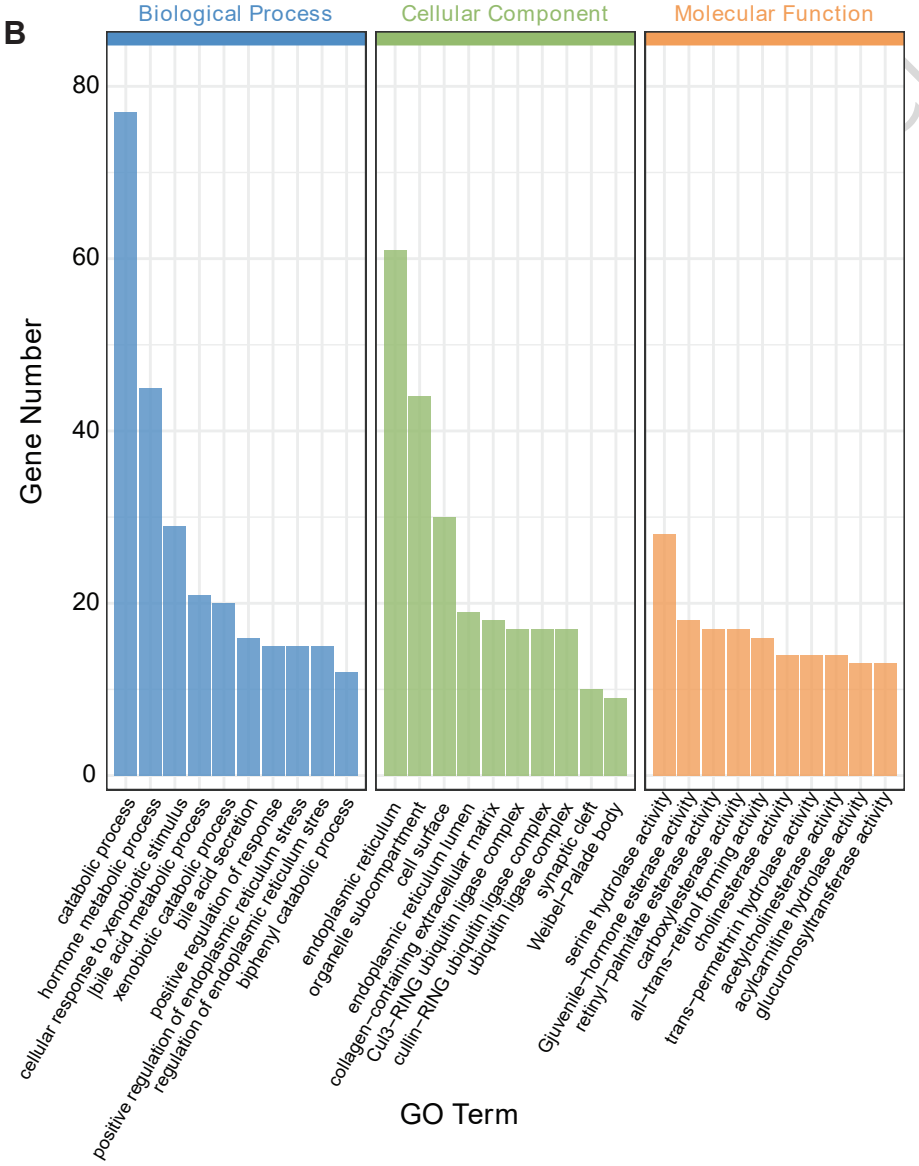
B

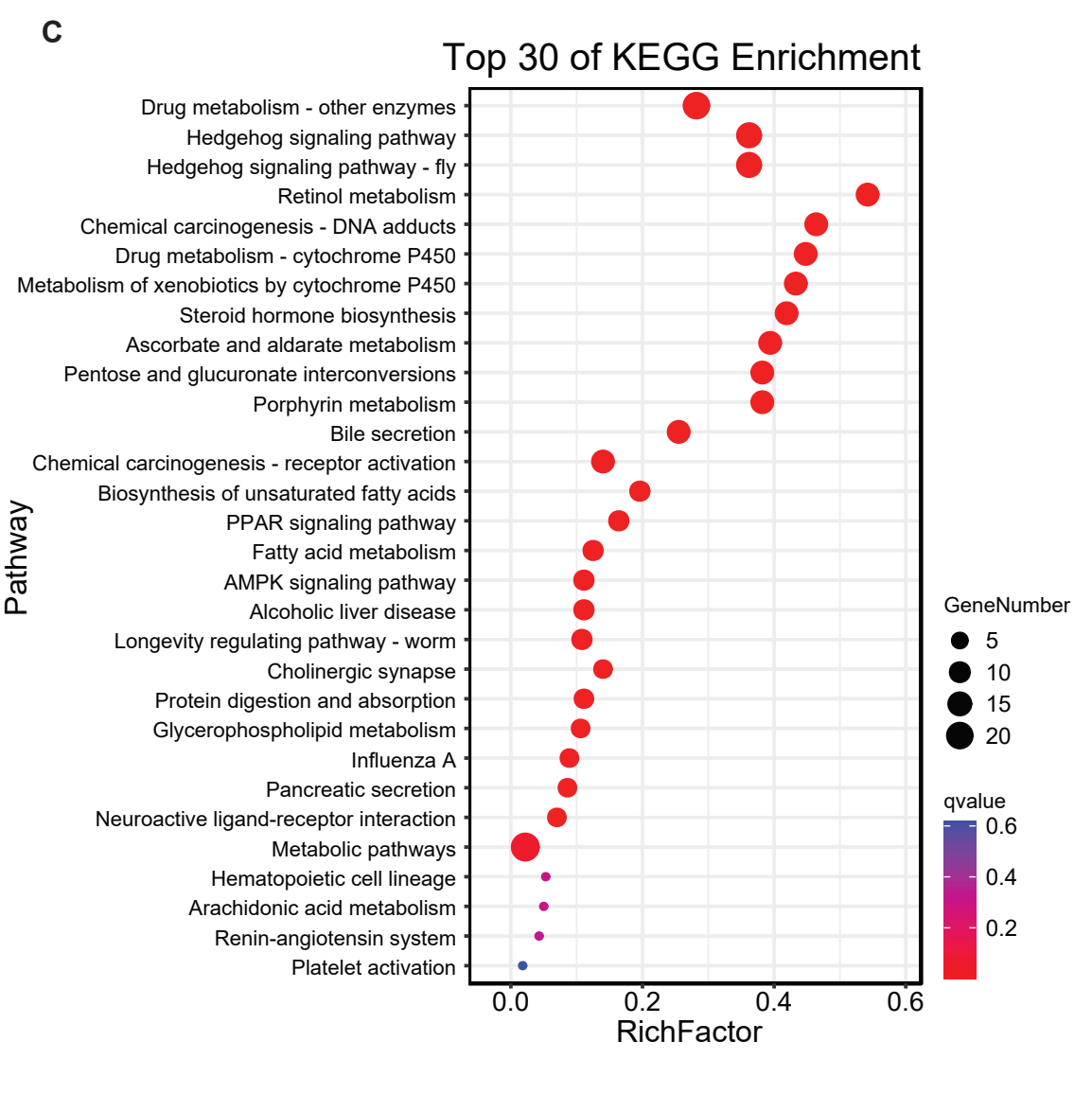
```
WT   TTATCCTTACCTCTATCAGAGGCCACC (94,593bp) CCTAGACCGACACCACCTCAACC
M1   TTATCCTTATTATCTATCAGAGGCCACC (94,593bp) CCTAGAC----ACCACCTCAACC (Δ-6bp+3bp)
M2   -------------------------- (94,577bp) CCTAGACCGACACCACCTCAACC (Δ-42bp)
M3   TTATCCTTACCTCTATCAGAGGCCACC (94,593bp) CCTAGAC----ACCACCTCAACC (Δ-4bp)
M4   TTATCCT-------ATCAGAGGCCACC (94,593bp) CCTAGACCGACACCACCTCAACC (Δ-7bp)
M5   TTATCCTTATC-CTATCAGAGGCCACC (94,593bp) CCTAGACC------ACCTCAACC (Δ-9bp+2bp)
```

C




A
+32/-909
+16/-770
+246/-487
+26/-460
+33/-2028
+186/-951
+25/-28
+443/-227
+289/-828
+154/-1523
+3/-87
+36/-627
+469/-206
+182/-376
+0/-0
+50/-1374
+65/-1563
Apis mellifera +110/-824
Megachile rotundata +214/-795
Harpegnathos saltator +311/-911
Solenopsis invicta +604/-690
Polistes fuscatus +553/-284
Vespa mandarinia +242/-658
Gonatopus flavifemur +910/-2781
Nasonia vitripennis +315/-1259
Pteromalus puparum +928/-683
Copidosoma floridanum +589/-2154
Belonocnema treatae +643/-2390
Hyposoter didymator +604/-660
Venturia canescens +856/-451
Gregopimpla kuwanae +327/-2124
Lysiphlebus fabarum +1517/-772
Microplitis demolitor +425/-1486
Orussus abietinus +171/-4930
Cephus cinctus +338/-4347
Athalia rosae +351/-4465
Apoidea
Formicidea
Vespoidea
Chrysidoidea
Chalcidoidea
Cynipoidea
Ichneumonoidea
Orussoidea
Cephoidea
Tenthredinoidea
Permian
Triassic
Jurassic
Cretaceous
Paleogene
Ng
-250
-200
-150
-100
-50
0
5000
10000
15000
20000
25000
single-copy
multi-copy
species-specific
unassign
other


B
Biological Process
Cellular Component
Molecular Function
Gene Number
0
20
40
60
80
catabolic process
hormone metabolic process
cellular response to xenobiotic stimulus
lbile acid metabolic process
xenobiotic catabolic process
bile acid secretion
positive regulation of response
positive regulation of endoplasmic reticulum stress
regulation of endoplasmic reticulum stres
biphenyl catabolic process
endoplasmic reticulum
organelle subcompartment
cell surface
endoplasmic reticulum lumen
collagen-containing extracellular matrix
Cul3-RING ubiquitin ligase complex
cullin-RING ubiquitin ligase complex
ubiquitin ligase complex
synaptic cleft
Weibel-Palade body
serine hydrolase activity
Glyuvenile-hormone esterase activity
retinyl-palmitate esterase activity
carboxylesterase activity
all-trans-retinol forming activity
cholinesterase activity
trans-permethrin hydrolase activity
acetylcholinesterase activity
acylcarnitine hydrolase activity
glucuronosyltransferase activity
GO Term


C
Top 30 of KEGG Enrichment
Drug metabolism - other enzymes
Hedgehog signaling pathway
Hedgehog signaling pathway - fly
Retinol metabolism
Chemical carcinogenesis - DNA adducts
Drug metabolism - cytochrome P450
Metabolism of xenobiotics by cytochrome P450
Steroid hormone biosynthesis
Ascorbate and aldarate metabolism
Pentose and glucuronate interconversions
Porphyrin metabolism
Bile secretion
Chemical carcinogenesis - receptor activation
Biosynthesis of unsaturated fatty acids
PPAR signaling pathway
Fatty acid metabolism
AMPK signaling pathway
Alcoholic liver disease
Longevity regulating pathway - worm
Cholinergic synapse
Protein digestion and absorption
Glycerophospholipid metabolism
Influenza A
Pancreatic secretion
Neuroactive ligand-receptor interaction
Metabolic pathways
Hematopoietic cell lineage
Arachidonic acid metabolism
Renin-angiotensin system
Platelet activation
Pathway
0.0
0.2
0.4
0.6
RichFactor
GeneNumber
5
10
15
20
qvalue
0.6
0.4
0.2