

RESEARCH ARTICLE

Open Access



An expanded odorant-binding protein mediates host cue detection in the parasitic wasp *Baryscapus dioryctriae* basis of the chromosome-level genome assembly analysis

Xiaoyan Zhu^{1,2,3}, Yi Yang^{1,2,3}, Qiuyao Li^{1,2,3}, Jing Li⁴, Lin Du^{1,2,3}, Yanhan Zhou^{1,2,3}, Hongbo Jin^{1,2,3}, Liwen Song⁴, Qi Chen^{1,2,3,5*} and Bingzhong Ren^{1,2,3,5*}

Abstract

Background *Baryscapus dioryctriae* (Chalcidodea: Eulophidae) is a parasitic wasp that parasitizes the pupae of many Pyralidae members and has been used as a biological control agent against *Dioryctria* pests of pinecones.

Results This *B. dioryctriae* assembly has a genome size of 485.5 Mb with a contig N50 of 2.17 Mb, and scaffolds were assembled onto six chromosomes using Hi-C analysis, significantly increasing the scaffold N50 to 91.17 Mb, with more than 96.13% of the assembled bases located on chromosomes, and an analysis revealed that 94.73% of the BUSCO gene set. A total of 54.82% (279.27 Mb) of the assembly was composed of repetitive sequences and 24,778 protein-coding genes were identified. Comparative genomic analysis demonstrated that the chemosensory perception, genetic material synthesis, and immune response pathways were primarily enriched in the expanded genes. Moreover, the functional characteristics of an odorant-binding protein (BdioOBP45) with ovipositor-biased expression identified from the expanded olfactory gene families were investigated by the fluorescence competitive binding and RNAi assays, revealing that BdioOBP45 primarily binds to the *D. abietella*-induced volatile compounds, suggesting that this expanded OBP is likely involved in locating female wasp hosts and highlighting a direction for future research.

Conclusions Taken together, this work not only provides new genomic sequences for the Hymenoptera systematics, but also the high-quality chromosome-level genome of *B. dioryctriae* offers a valuable foundation for studying the molecular, evolutionary, and parasitic processes of parasitic wasps.

Keywords Hi-C, Parasitic wasp, *Baryscapus dioryctriae*, Odorant-binding protein

*Correspondence:

Qi Chen
chenq766@nenu.edu.cn
Bingzhong Ren
bzren@nenu.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

Parasitic wasps are in the spotlight for their potential as biological control agents against pests and for their megadiverse biology, positioning them as the subject of biological research [1]. Various parasitic wasps' species have been widely used as effective natural enemies against numerous destructive pests in agriculture and forestry. For example, *Chouioia cunea* is used for controlling *Hyphantria cunea* and *Helicoverpa armigera* [2], and *Trichogramma brassicae* is used for controlling *Ostrinia nubilalis* [3]. Given the diverse biological characteristics of parasitic wasps, they are gradually becoming an ideal model study system for investigating insect genetic diversity [4, 5]. The number of haploid chromosomes in parasitic wasps ranges from 3 to 23 [6]. With the development of new sequencing technologies, numerous insect genome assemblies have now been improved to the chromosome level [7]. To date, data for 14 genome assemblies of parasitic wasps at the chromosome level have been published, and 10 of these assemblies rely on high-through chromosome conformation capture (Hi-C) technology, which include those of *Aphidius gifuensis* (Hi-C, $n=6$) [8], *Aphelinus certus* ($n=4$), *Aphelinus atriplicis* ($n=4$) [9], *Cotesia glomerata* (Hi-C, $n=10$) [10], *Cotesia congregata* (Hi-C, $n=10$) [11], *Chelonus formosanus* (Hi-C, $n=7$) [12], *C. cunea* (Hi-C, $n=6$) [13], *Eretmocerus hayati* (Hi-C, $n=4$) [14], *Microplitis manila* (Hi-C, $n=11$) [15], *Nasonia vitripennis* ($n=5$) [16], *Pteromalus puparum* (Hi-C, $n=5$) [1], *Telenomus remus* (Hi-C, $n=10$) [17], *Theocolax elegans* (Hi-C, $n=7$) [18], and *Venturia canescens* ($n=11$) [19]. This approach provides additional and essential information that is restricted by fragmented genomes and improves comparative genome analyses of parasitic wasps, which is helpful for further analysis of parasitoid lifestyle, behavior, and living habits [7]. Comparative genomics studies have revealed that the oxyglycosylation of the mucin domain in the hemomucin protein plays a pivotal role in the passive immune evasion mechanism employed by parasitoids [20]. Furthermore, a cluster of genes displaying remarkably accelerated evolutionary rates has been identified in these two miniaturized parasitoids, potentially attributable to their convergent adaptation toward miniaturization [17].

The antennae and ovipositors of parasitic wasps have long been linked to the identification of chemical cues pertaining to host localization, inspection, and oviposition [21]. Parasitic wasps antennae have been extensively studied, including ultrastructural studies of sensilla [22], electrophysiological detection of olfactory-active volatiles [23], identification of olfactory-related genes through transcriptome sequencing [24], and functional characterization of the olfactory genes expressed in the antennae [22–26]. Although the sensilla of some parasitic wasps

have been successfully identified in the ovipositor [27], and some olfactory-related genes have been detected in the ovipositor; however, the functions of these genes have yet to be characterized [28]. For example, in the fig wasp *Apocrypta westwoodi*, their ovipositors not only serve as egg-laying organs but also function as olfactory organs, responding to volatile compounds and CO₂ in a gaseous form [29]. The olfactory function of insect ovipositors is supported by greater data in other species, and investigations have shown that the ovipositors of several Lepidoptera species express genes encoding olfactory receptors [30]. In addition, through morphological studies, olfactory sensilla have been shown to be present in the ovipositors of *Monopis crocipitella* and *Homoeosoma nebulella* through morphological studies [31, 32]. Functional characterization of neural responses in the sensors on *Manduca sexta* ovipositors via single sensillum recordings revealed that these sensilla indeed house functional olfactory sensory neurons (OSNs) [33]. The highly expressed odorant receptors (ORs) in the ovipositors were subsequently confirmed to be involved in the detection of host plant volatiles in *Helicoverpa assulta* [34]. In Diptera, *Bactrocera dorsalis* was also found to recognize ovipositional preferences by ovipositors [35].

The insect olfactory system primarily encompasses olfactory proteins, generally odorant-binding proteins (OBPs), chemosensory proteins (CSPs), sensory neuron membrane proteins (SNMPs), Niemann-Pick type C2 proteins (NPC2s), ionotropic receptors (IRs), gustatory receptors (GRs) and odorant receptors (ORs) [24]. OBPs are small soluble proteins that are enriched in sensillum lymph; they act as semio-chemical carriers by binding and transporting these chemicals to ORs, and the process mediated by OBPs is considered the first stage of insect olfactory perception [36, 37]. In parasitic wasps, CcunOBP2 in *C. cunea* and MmedOBP14 in *Microplitis mediator* participate in the detection of plant volatiles, which may contribute to locating habitats, supplementing nutrients and searching for plant hosts [36, 38]. In addition to their main chemical senses, insect OBPs are also present in other peripheral appendages and internal organs and participate in a variety of insect life activities, such as temperature and humidity perception, mating, taste, and development [39]. OBPs are also expressed within female insect ovipositors, and histological analysis has revealed the expression of OBPs in mosquito ovaries and eggshells [40–42]. Furthermore, the female ovary of the oriental fruit fly *B. dorsalis* has significant expression of OBP19c [43]. OBPs were also shown to be abundantly expressed in the ovipositor of parasitic wasps [28]. Regarding the function of OBPs in Diptera, the female abdomen-biased expression of OBP56d-1 in *B. dorsalis*

significantly increased after stimulation by the oviposition inducer 1-octen-3-ol [44], and the highly expressed OBP56d and OBP56d-2 in the ovipositor were found to be jointly responsible for the oviposition preference of 3-hexenyl acetate (3-HA) [35]. This finding suggests that OBPs on the ovipositor may serve to bring odors or pheromones into proximity to the odor receptors present in the female reproductive tract. Therefore, characterizing the specific OBPs expressed in the ovipositor will help establish a connection between OBPs and their potential functions.

Baryscapus dioryctriae (Chalcidoidea: Eulophida; Fig. 1) is a gregarious endoparasitoid wasp of many Pyralidae pest pupae [45]. In view of the two important problems in the application of parasitic wasps, reproductive efficiency and field application efficiency, and our previous research revealed that 2-butyl-2-octenal could significantly improve the reproductive efficiency of *B. dioryctriae* during studying the olfactory function of NPC2s in this wasp [46]. The control efficiency of parasitizing wasps in field applications largely depends on their host-searching ability, which usually relies on their responses to the semiochemicals related to their hosts, and such responses may be influenced by genetic mechanisms and learning behaviors [26, 47]. Parasitic wasps can be trained through learning behaviors to increase their efficiency in locating hosts in laboratory settings or in the field [48, 49]. Therefore, the identification of active compounds associated with host search will be beneficial for the future development of appropriate parasitoid

olfactory conditioning methods for the optimization of biological pest control.

In our previous transcriptomic analysis of *B. dioryctriae*, six OBP genes were found to be more highly expressed in the ovipositor than in other tissues [28]. To obtain more complete genetic information, in our ongoing project, a high-quality chromosome-level genome assembly of *B. dioryctriae* was generated using a combination of Illumina, PacBio, and Hi-C technologies. A comparative genomic analysis was conducted, leading to the re-identification for odorant-binding proteins of *B. dioryctriae*' odorant-binding proteins (*BdioOBPs*) based on a high-quality genome assembly. Within the expanded OBP genes, a gene exhibiting partial ovipositor expression was discovered and subsequently characterized for its function. Comparative genomic analysis and identification of *BdioOBPs* were performed based on high-quality genomic assembly. Highly specific expression of the expanded OBP gene was detected in the ovipositor, and then its function was deeply characterized in vitro and vivo.

Results

Chromosome-level genome assembly of *B. dioryctriae*

K-mer distribution analysis (K=19) predicted that *B. dioryctriae* had an estimated genome size of 486.86 Mb with a low heterozygosity rate of 0.39% and a modest GC of 38.18% (Additional file 1: Fig. S1), which indicated that *B. dioryctriae* has a simple genome and facilitated genome assembly. Single-molecule real-time PacBio reads (10 Gb)

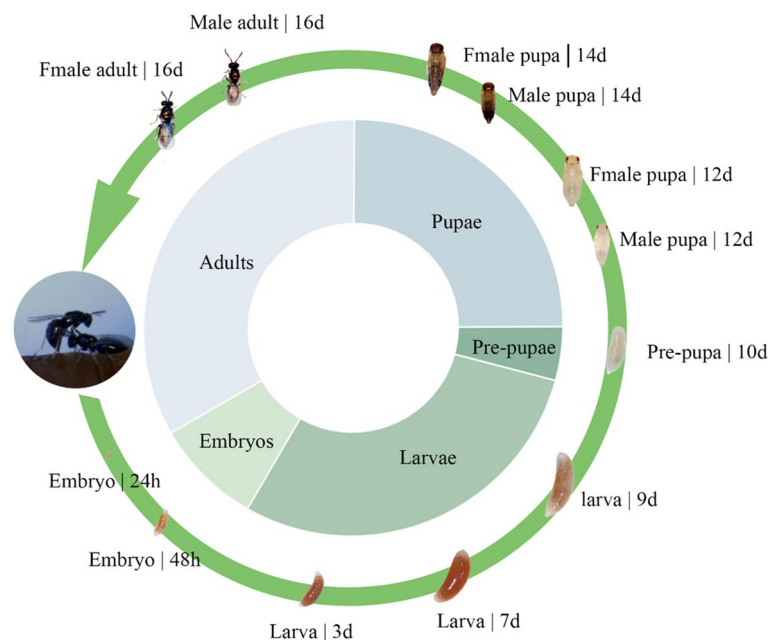


Fig. 1 Life cycle of the parasitoid wasp *B. dioryctriae* on its alternative host *Galleria mellonella*. Female wasps lay eggs in the host chrysalis

were used (Additional file 2: Table S1). We first generated a contig-level assembly with a genome size of 509.39 Mb, a contig N50 length of 2.17 Mb, and a GC content of 38.99% for the genome of *B. dioryctriae* (Table 1). Hi-C data were subsequently used to improve the assembly, with more than 96.13% of the assembled sequences anchored to six chromosomes; this outcome was further supported by karyotype analysis (Fig. 2A and B). The chromosomal assembly of *B. dioryctriae* spans 485.5 Mb with a scaffold N50 length of 91.17 Mb, and the length of the largest chromosome was 109.30 Mb (Table 1, Additional file 1: Fig. S2). More than 98.64% of Illumina reads were successfully mapped to the assembly, indicating its excellent integrity and accuracy. Moreover, genome assessment using benchmarking universal single-copy orthologs (BUSCO) indicated that 94.73% of the insect gene sets were present and complete.

Genome annotation and genomic characteristics

A total of 54.82% (279.27 Mb) of the genome of *B. dioryctriae* was composed of repeat sequences. Transposable elements (TEs) accounted for 48.24% of the genome of *B. dioryctriae*, of which retroelements were the most abundant TE group (30.41%, 154.90 Mb), followed by DNA transposons (17.84%, 90.86 Mb) (Additional file 2: Table S2). Annotation revealed 33.5 Mb of tandem repeats, which accounted for 6.58% of the whole genome (Additional file 2: Table S2). Genome annotation revealed

that 32.02% (155.45 Mb) of the genome sequences encoded proteins in *B. dioryctriae*, with a total of 24,778 protein-coding genes identified. Notably, this gene count is comparable to those observed in two species characterized by large genomes: *G. flavifemur* (23,056) and *E. hayati* (23,911). The average gene length was determined to be 6273.86 bp, while the average coding sequence (CDS) length was 1447.44 bp. The average exon length and average intron length were 1631.1 bp and 96,160 bp, respectively; these values were consistent with those reported for most species except for *A. mellifera* and *N. vitripennis*, which presented high intron ratios (Additional file 2: Table S3 and 4). Additionally, approximately 90.48% of the genes (22,419 genes) were functionally annotated (Additional file 2: Table S5). We also identified 95.1% of the BUSCO Insecta database (Insecta_odb9) genes at the protein level.

The analysis of chromosome synteny was carried out by conducting a pairwise synteny search of CDS using the MCScanX pipeline. The results indicate that numerous genome rearrangements have taken place since the split of *B. dioryctriae* and *C. cunea* (according to this study, approximately 53.7 million years ago), as shown in the synteny map (Fig. 2C). It can be observed from the collinearity map that Chromosome1 (Chr1) in *N. vitripennis* shows good collinearity with Chr3 and Chr6 in the Eulophidae wasp genomes (*B. dioryctriae* and *C. cunea*). *B. dioryctriae* and *N. vitripennis* exhibit lower levels of homology than *C. cunea*. In this analysis, 253 syntenic blocks were identified between *B. dioryctriae* and *C. cunea*, containing 5–411 genes, with an average of 38.27 genes per syntenic block (Additional file 2: Table S6). This comparison revealed 479 syntenic blocks between *B. dioryctriae* and *N. vitripennis*, containing anywhere from 5 to 81, with an average of 16.42 genes per syntenic block (Additional file 2: Table S7).

Table 1 Genome assembly of *B. dioryctriae*

	Type	Contig length	Contig number	
Primary genome assembly	N50	2,170,212		
	N90	426,329		
	Longest	11,935,700		
	Total	509,389,314	876	
	Length > 100 bp	509,428,314	876	
	Length > 2 kb	509,428,314	876	
	GC (%)	38.99		
	Type	Scaffold length	Contig number	Scaffold number
Chromosome-level genome assembly	LG01 (Chr1)	109,293,743	85	
	LG02 (Chr2)	66,102,127	55	
	LG03 (Chr3)	91,163,347	79	
	LG04 (Chr4)	102,631,559	70	
	LG05 (Chr5)	90,521,596	89	
	LG06 (Chr6)	25,802,483	18	
	Total	485,514,855	396	
	N50	91,171,147		3
	N90	66,107,527		5
	Longest	109,302,143		

Orthology and phylogenetic relationships

A total of 22,458 orthogonal groups were identified among *B. dioryctriae* and 13 other Hymenopteran species. The gene family clusters were divided into five categories: generic single-copy genes, multicopy genes, species-specific multicopy genes, species-specific single-copy genes, and other genes. There were 920 species-specific orthogroups and 3939 species-specific genes identified in the genome of *B. dioryctriae* (Additional file 2: Table S8).

The phylogenetic relationships between *B. dioryctriae* and the other Hymenopteran insects used in this study were determined with a genome-wide set of 1265 single-copy genes. As expected, the maximum likelihood (ML) phylogenetic analysis demonstrated the separation of the two Symphytas (*O. abietinus* and *A. rosae*) from the

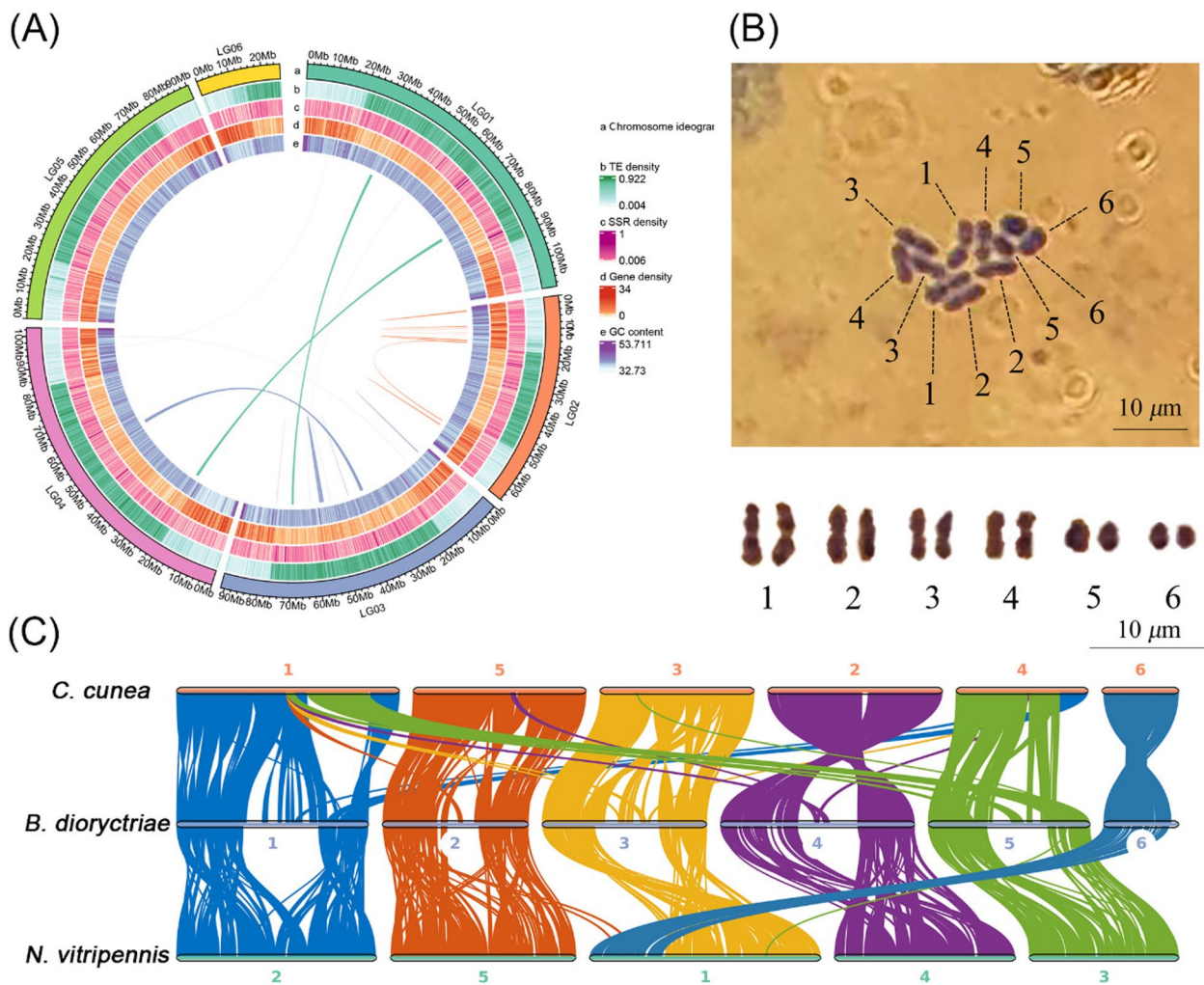


Fig. 2 Comprehensive genomic insights and chromosomal synteny of *B. dioryctriae* in comparison with *C. cunea* and *N. vitripennis*. **A** Genome landscape of the parasitoid wasp *B. dioryctriae*. From the outer to inner circles: (a) six chromosomes at the Mb scale; (b, c, and d) TE density, SSR density, and gene density across the genome; (e) GC contents across the genome. **B** Karyograms of *B. dioryctriae*. **C** Chromosome synteny based on CDS pairwise alignment between *B. dioryctriae*, *C. cunea*, and *N. vitripennis*. The colored lines indicate shared syntenic blocks

12 Apocritas, consistent with expectations. The eight parasitic wasps (*B. dioryctriae*, *C. cunea*, *N. vitripennis*, *P. puparum*, *E. hayati*, *T. pretiosum*, *T. remus*, and *A. gifuensis*) clustered together, with four Aculeatas (*A. mellifera*, *O. biroi*, *G. flavifemur* and *Polistes dominula*) serving as a sister group. The six chalcidoids (*B. dioryctriae*, *C. cunea*, *N. vitripennis*, *P. puparum*, *T. pretiosum* and *E. hayati*) clustered together, and the Eulophidae wasps (*B. dioryctriae* and *C. cunea*) were more closely related to the two Pteromalidae (*P. puparum* and *N. vitripennis*) than were Trichogrammatidae (*T. pretiosum*) and Aphelinidae (*E. hayati*) (Fig. 3A). *B. dioryctriae* and other Chalcidoidea species are estimated to have diverged from Cretaceous with a divergence time of approximately 53.71 million years.

Genes under expansion and contraction

The gene family evolution analysis conducted using CAFE identified 309 expanded gene families and two contracted gene families on the terminal branch of *B. dioryctriae* was separated from other parasitic wasps (Fig. 3A). Notably, several gene families associated with chemosensory activity, including OR, IR and GR were found to be expanded in the genome of *B. dioryctriae* (Additional file 2: Table S9 and S10). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses revealed that the terms in which *B. dioryctriae* expanded gene families were enriched were odorant binding (GO: 0005549, 170 genes, $P=7.38E-33$), nucleotide binding (GO: 0000166, 152 genes, $P=3.40E-51$), DNA-directed DNA polymerase activity (GO: 0003887,

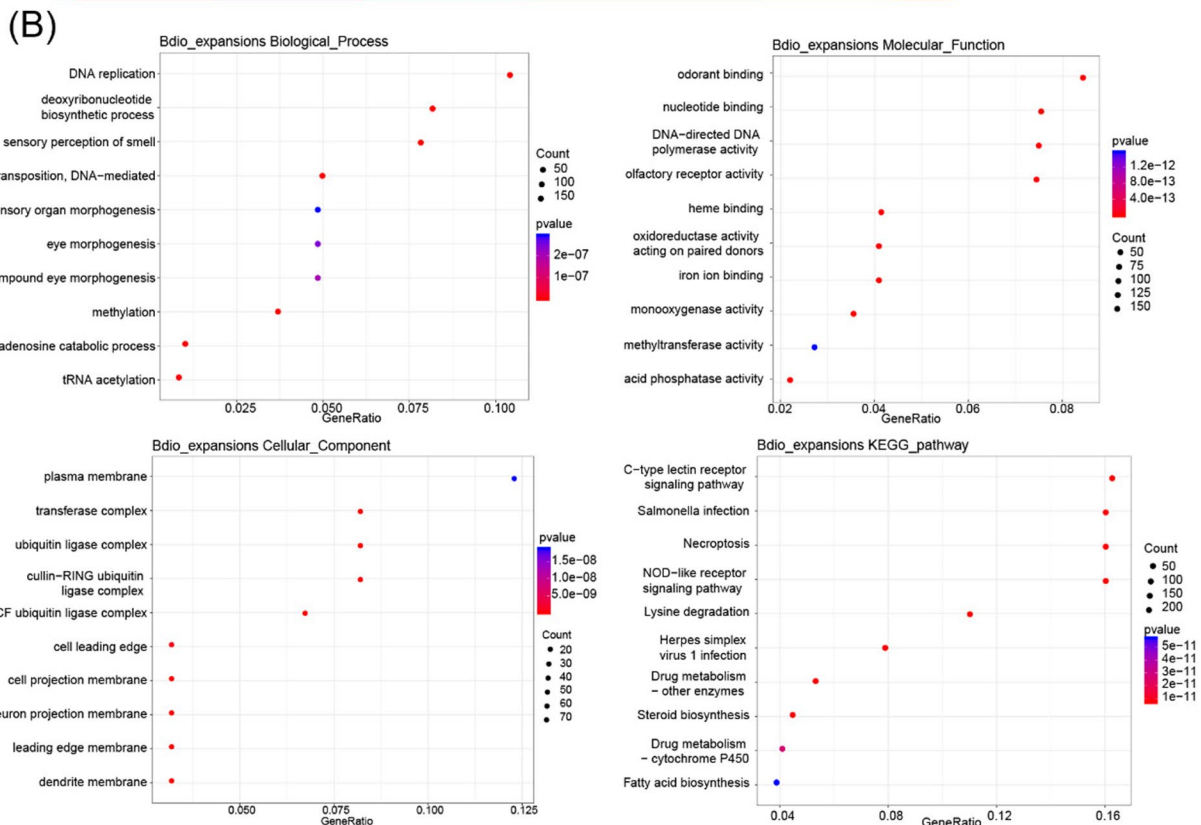
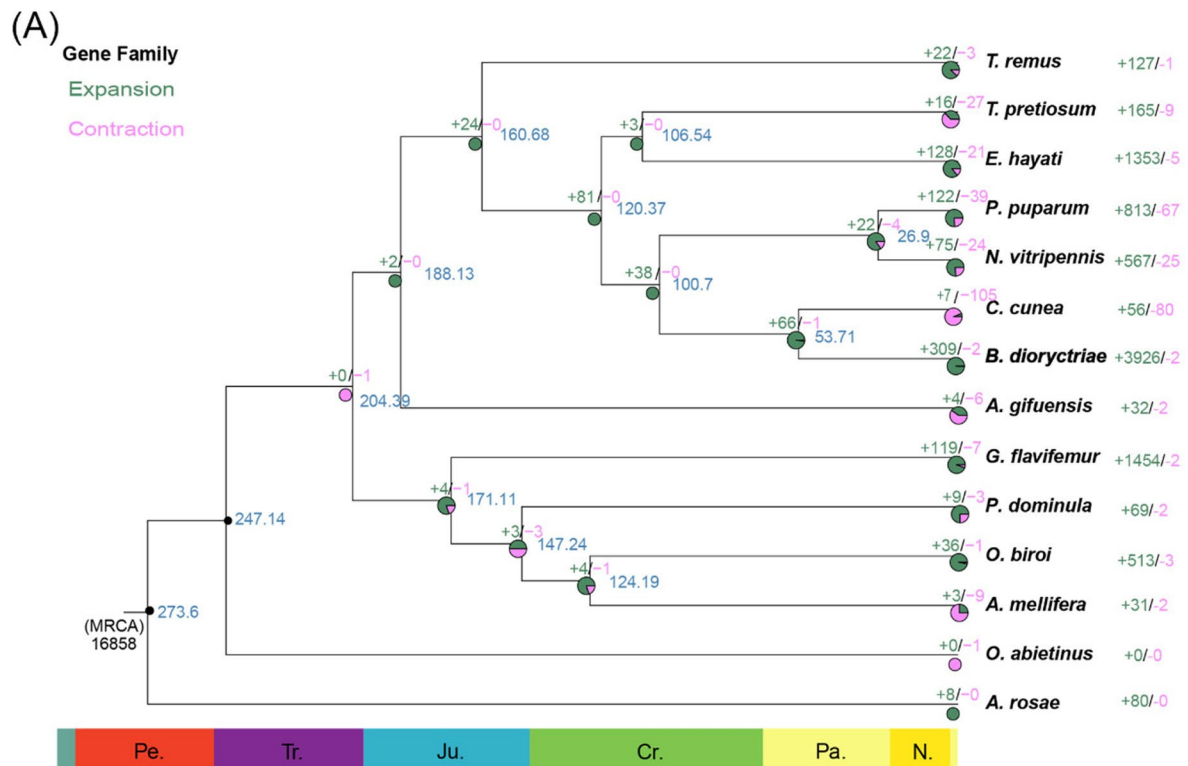


Fig. 3 Comparative genomic analysis of the parasitoid wasp *B. dioryctriae*. **A** A maximum-likelihood phylogenetic tree is shown for *B. dioryctriae* with 13 hymenopteran insects. The phylogenetic tree was based on 1265 single-copy proteins. *A. rosae* was used as the outgroup. The bootstrap value of all nodes is supported at 100/100. The numbers of expanded (green) and contracted (pink) gene families are shown on the branches. **B** The GO and KEGG enrichment results of *B. dioryctriae* difficile-expanded genes are summarized and visualized as scatter plots

151 genes, $P=6.32E-66$), olfactory receptor activity (GO: 0004984, $P=7.48E-32$), DNA replication (GO: 0006260, 153 genes, $P=1.39E-40$), sensory perception of smell (GO: 0007608, 115 genes, $P=1.61E-17$), deoxyribonucleotide biosynthetic process (GO: 0009263, 120 genes, $P=1.2E-15$), C-type lectin receptor (CLR) signaling pathway (ko04625, 207 genes, $P=3.7E-63$), NOD-like receptor (NLR) signaling pathway (ko04621, 204 genes, $P=9.26E-65$), necroptosis (204 genes, $P=1.75E-64$), and Salmonella infection (ko05132, 204 genes, $P=4.05E-63$). Only two gene families contracted, GO and KEGG enrichment analyses of only two gene families revealed that the terms in which *B. dioryctriae* contracted gene family was enriched were serine-type endopeptidase activity (GO: 0004252, one gene, $P=0.02$) and neuroactive ligand–receptor interaction (ko04080, one gene, $P=0.04$) (Fig. 3B).

Functional analysis of BdioOBP45

A total of 81 OBPs were identified from the genome of *B. dioryctriae* (Additional file 2: Table S11). All 81 OBP genes were mapped onto six chromosomes (Fig. 4A) and mainly clustered on Chr4 (64 genes, 74.70%), Chr5 (11 genes, 13.25%) and Chr6 (5 genes, 6.02%) (Fig. 4A). Two OBP gene families (OG0000305 and OG0006966) were found to be expanding (Additional file 2: Table S9). The OBP genes were distributed on Chr4 and Chr6; the expression profiles of these OBP genes were relatively scattered; *BdioOBP45* was highly expressed specifically on the ovipositor, with a fragments per kilobase million (FPKM) value of 8804.74 (Fig. 4B); and the ovipositor-biased expression pattern of *BdioOBP45* was validated by real-time quantitative polymerase chain reaction (RT-qPCR) (Additional file 1: Fig. S3).

The recombinant protein BdioOBP45 was expressed and purified in a bacterial system (Additional file 1: Fig. S4). A binding assay showed that BdioOBP45 could bind efficiently to 1-N-phenyl-naphthylamine (1-NPN) with a dissociation constant (K_d) of 2.324 μ M (Fig. 5A). In addition, saturation and linear Scatchard plots show a single binding site for BdioOBP45 and 1-NPN. To investigate the function of BdioOBP45 in odor recognition, gas chromatography-mass spectrometry (GC–MS) analysis was conducted, revealing the presence of a total of 47 compounds in *P. koraiensis*

cones, pest-damaged *P. koraiensis* cones, mechanically damaged *P. koraiensis* oleoresin, pest-damaged *P. koraiensis* oleoresin, and *D. abietella* pupae (Additional file 1: Fig. S5, Additional file 2: Table S12). Of these, 55 compounds were successfully acquired for subsequent binding assays. Four of the 55 tested compounds (Additional file 2: Table S13) reduced the relative fluorescence intensity of the BdioOBP45/1-NPN complex by > 50%. Three of them strongly bound to BdioOBP45 ($K_i < 10 \mu$ M), 7.52 \pm 0.55 μ M caryophyllene oxide, 9.59 \pm 0.27 μ M g-terpinene and 9.7 \pm 0.53 μ M α -pinene (Table 2, Fig. 5B).

RNA interference and behavior assays

To further investigate the biological functions of ligands that may have a certain binding affinity for BdioOBP45, we conducted Y-tube olfactory assays. The results showed that α -pinene and g-terpinene could significantly attract *B. dioryctriae* (Fig. 6A and B). Subsequently, to test the effect of reducing the transcription level of the BdioOBP45 gene on the behavior of *B. dioryctriae*, dsRNA was injected into female wasps. The optimal timing for behavioral detection was determined by RT-qPCR, and the attractiveness to α -pinene and g-terpinene was compared. Initially, we compared the transcription levels of BdioOBP45 in the control group (non-injected), dsGFP-injected group, and dsBdioOBP45-injected wasps at 12 h, 24 h, 36 h, and 48 h post-injection. Throughout the study period, there was no difference in the transcription level of BdioOBP45 in the control group (non-injected) and the dsGFP-injected group (Fig. 6C). The transcription level of *BdioOBP45* was significantly reduced at all time points ($P < 0.05$), with the greatest reduction observed 24 h post-injection, which was 98% lower than the baseline level. Then, we compared the behavioral changes in *B. dioryctriae* between different treatment groups 24 h after dsRNA injection, the Y-tube olfactometer test results revealed that the attractiveness to g-terpinene was significantly lower in the dsBdioOBP45-injected group than in the non-injected and dsGFP-injected groups, whereas no significant difference was found between the two control groups. There was no significant change in the response to α -pinene after injection (Fig. 6D).

(See figure on next page.)

Fig. 4 Genome-wide analysis of odorant-binding proteins and gene expression patterns in *B. dioryctriae* with a focus on OG0000305 and OG000696. **A** Odorant-binding proteins in the *B. dioryctriae* genome. **B** Gene expression of OG0000305. **C** Gene expression of OG000696. FA and MA: female and male antennae; FH and MH: female and male heads without antennae; Fab: female abdomens without ovipositors and digestive tracts; Fov: female ovipositors; Mge: male genitalia; Mab: male abdomens without genitalia and digestive tracts; T: male and female thoraxes; L: male and female legs

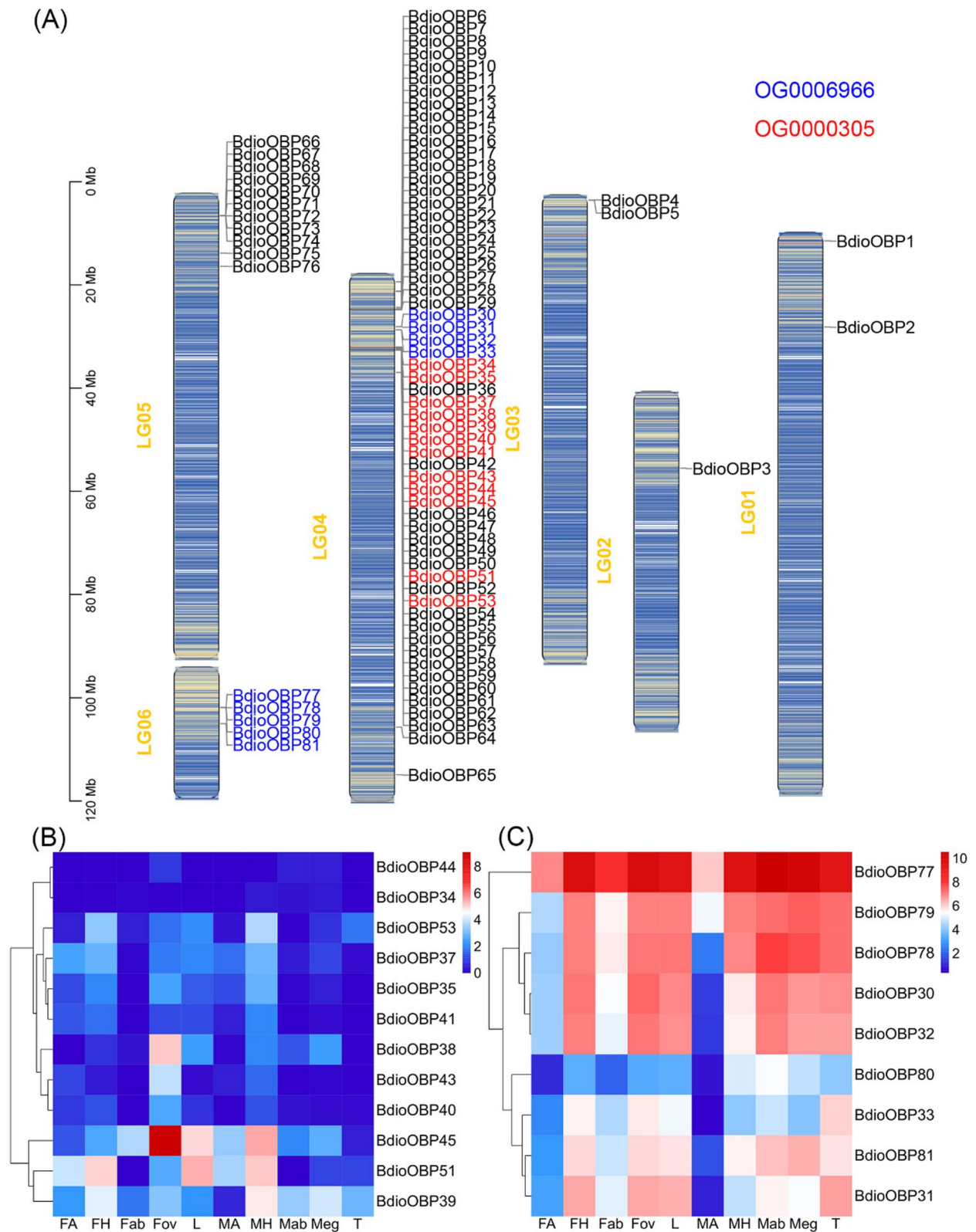


Fig. 4 (See legend on previous page.)

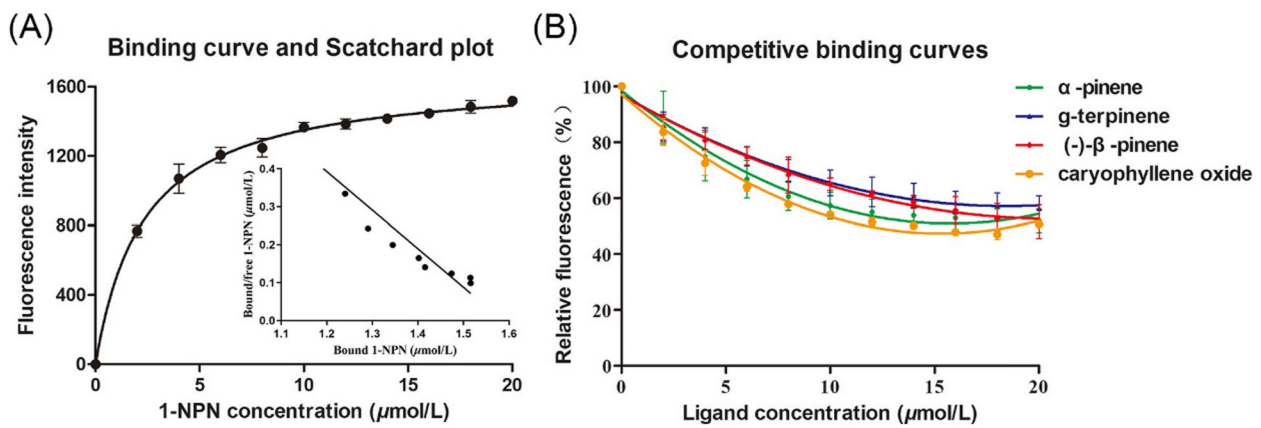


Fig. 5 Binding analysis of recombinant BdioOBP45 to candidate ligands. **A** The binding curve of 1-NPN to recombinant BdioOBP45 and the Scatchard plot. **B** Binding curves for BdioOBP45 binding to four ligands

Table 2 Binding affinities of BdioOBP45 to tested compounds

Ligands	IC50 (μM)	Ki (μM)
Caryophyllene oxide	13.98 \pm 1.03	7.52 \pm 0.55
α -Pinene	17.48 \pm 3.63	9.4 \pm 1.95
g-Terpinene	19.32 \pm 4.26	10.38 \pm 2.29

Discussion

Based on current sequencing projects, Hymenoptera genomes are modest in size (80% are between 180 and 340 Mb), although there are some exceptions; for example, *G. flavifemur* (636.5 Mb) and *E. hayati* (692.1 Mb), both of which are relatively large genomes in Hymenoptera [14, 50], were selected for the comparative genome

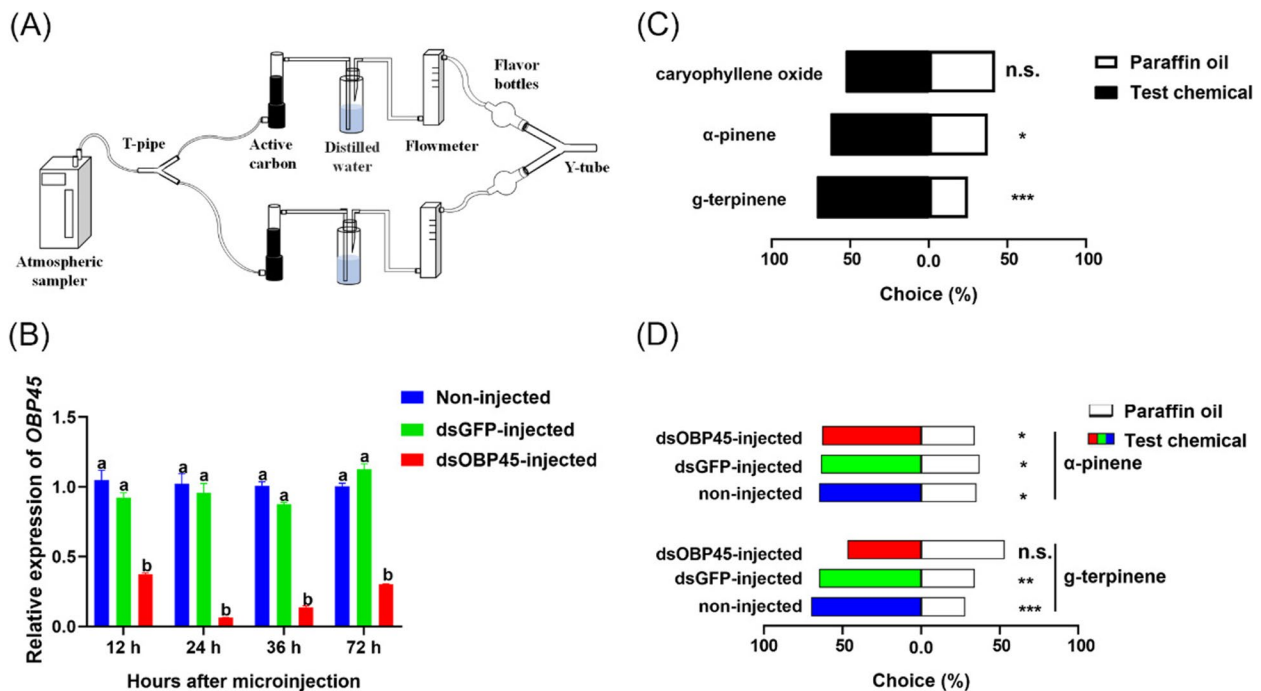


Fig. 6 In vivo functional investigation of BdioOBP45. **A** Behavioral experiment device diagram for *B. dioryctriae*. **B** Behavioral responses of *B. dioryctriae* adults to three ligands of BdioOBP45. **C** Relative expression levels of *BdioOBP45* at 12 h, 24 h, 36 h, and 48 h after microinjection of double-stranded RNA (dsRNA) into wasps. Asterisks on the top portion of the bars show that the values differed significantly; that is, columns labeled with different letters are significantly different ($P < 0.05$). **D** Behavioral response of *B. dioryctriae* BdioOBP45-silenced adult wasps to g-terpinene and α -pinene

in this study. The expansion of repetitive sequences is one of the most important factors for increasing genome size, and this phenomenon has been reported in many insects [50]. This also applies to Hymenopteran insects, such as *E. hayat*, with a repeat sequence percentage of 370.8 Mb, and *G. flavifemur*, with a repeat sequence percentage of 381.7 Mb. The repeat sequences of *B. dioryctriae* accounted for 54.82% (279.27 Mb) of the whole genome, and the average length of each gene and intron length were moderate, indicating that the TE sequences were amplified. In particular, the amplification of retroelement sequences was a major contributor to the genome enlargement of *B. dioryctriae*. TE amplification and insertion can cause various changes in the host genome, such as chromosomal rearrangements, gene disruption, and regulation of gene expression; moreover, TEs are associated with a variety of insect adaptations, including resistance to insecticides [51], temperature adaptation [52], adaptive evolution of natural enemy defense [53], and synchronous evolution of phenotypic changes [54]. These retrotransposons may be related to *B. dioryctriae*'s adaptability to low temperatures and esticides.

Significant expansion or contraction of gene families is often associated with adaptive differentiation of species [55, 56]. For example, *Megastigmus duclouxiana* and *Megastigmus sabinae*, which are found in alpine habitats where solar UV radiation increases with altitude, have expanded gene families that are enriched in eye development [57]. The comparative genomic analysis revealed that the expanded genes in *B. dioryctriae* were primarily associated with chemosensory perception and genetic material synthesis in regions enriched with gene ontology terms. Insects rely on their chemosensory system to discriminate between many attractive or aversive chemical cues in the environment, which is essential for taking appropriate action in response [58]. Parasitic wasps use their sensitive olfactory system to discriminate host-related phytochemicals to determine host location and oviposition site selection [26]. There is a broad positive correlation between the complexity of the chemical ecology of arthropod species and the number of their chemosensory genes [59]. For example, the size of the wasp OR gene family reflects the complexity of chemical cues in its habitat [26]. The expansion of olfaction-related genes in *B. dioryctriae* might be related to the responsible forest environment in which it lived. KEGG annotation revealed that the expanded gene families were mainly enriched in the C-type lectin receptor signaling pathway, NOD-like receptor signaling pathway, and other immune-related pathways. Unlike vertebrates, which possess adaptive and innate immune responses, insects rely solely on the innate immune system to defend against microbes [60]. Pattern recognition receptors (PRRs) play an important

role in innate immunity. The NLR and CLR are six types of PRRS [61]. NLRs are intracellular molecules that recognize pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs) [62]. CLRs are transmembrane receptors that recognize carbohydrates and mediate immune responses [63]. The expansion of immune-related genes in *B. dioryctriae* suggests that *B. dioryctriae* has potential as part of an integrated pest management strategy (pest control combined with elimination of insect pathogens by its strong auto-immune system). Furthermore, a decreasing trend was observed in *Ceratosolen solmsi* for genes associated with chemical perception, detoxification metabolism, and various immune-related pathways, which may be related to its obligate parasitism [64]. Therefore, we suggest that the amplification of alleles and immune-related genes in *B. dioryctriae* may be related to its wide host range.

OBPs are crucial components in the olfactory system that mediate odor recognition, and they facilitate the recognition of odors by transporting semiochemicals to ORs through sensillum lymph, ultimately leading to the transmission of electrical signals to the insect brain. Additionally, OBPs play a role in protecting odorants from odorant-degrading enzymes [65]. Research indicates that the number of genes encoding OBPs varies significantly among parasitoid species, and numerous studies have verified that OBPs are vital for the host location of parasitoids through various methods such as gene silencing, fluorescence competitive binding assays, molecular docking, and so on [38]. The number of OBP genes in various parasitoid species ranges from 2 (*Scleroderma guani*) [66] to 98 (*N. vitripennis*) [67]. A total of 81 OBP genes were identified in *B. dioryctriae*, with the relatively high numbers potentially being associated with its complex living environment. Furthermore, the diverse counts of OBP genes in different species may be attributed to the depth of RNA sequencing [68]. In this study, a more comprehensive *B. dioryctriae* OBP gene family was identified based on genomics in comparison with the previously identified transcriptome-based results (27) [28]. A total of 81 OBP genes were primarily located on Chr4, Chr5, and Chr6. These genes showed a relatively clustered distribution, indicating that they may have undergone family expansion through tandem repeats. There are two expanded gene families among these genes. Through analysis of the expression levels of these two gene families, an interesting gene, *BdioOBP45*, was found to be highly expressed specifically in the ovipositor. The results of the fluorescence binding assay demonstrated that the main compounds that bind to *BdioOBP45* in vitro are caryophyllene oxide, α -pinene, and g-terpinene. All three compounds, which belong to the terpenoid class, are volatile terpenes that mediate interactions between plants

and insects, typically benefiting plants while harming herbivores [69]. According to the GC–MS results, all three compounds are typical pest-induced volatile compounds. Herbivore-induced plant volatiles (HIPVs) are important cues for natural enemies to find their hosts [70]. In subsequent behavioral experiments, both α -pinene and g-terpinene could elicit a significant behavioral attraction in *B. dioryctriae*. After *BdioOBP45* was silenced, the attraction of *B. dioryctriae* to g-terpinene was abolished, whereas the attraction to α -pinene remained, indicating that there may be other *BdioOBPs* in *B. dioryctriae* that can bind to α -pinene [37]. Furthermore, certain volatiles that do not inherently elicit a response in parasitoids can act as background volatiles to enhance the parasitoids' response to other volatiles [71]. Thus, caryophyllene oxide alone does not attract *B. dioryctriae* directly and may function synergistically in conjunction with other compounds to facilitate host location by parasitoids. To sum up, *BdioOBP45* could be used as a potential target to regulate the behavior of the natural enemy parasitic wasp *B. dioryctriae*.

Conclusions

In this study, a high-quality chromosome-level genome assembly of *B. dioryctriae* in parasitic wasps was generated using the Illumina, PacBio, and Hi-C combination strategy. Basic genomic analysis revealed the genomic characteristics, phylogenetic position, and gene family evolution of *B. dioryctriae*. Comparative genome analysis revealed many expanded genes in *B. dioryctriae* that were mainly involved in olfactory perception, genetic material synthesis, and immune response pathways. Eighty-one *BdioOBP* genes were identified and located in the genome. Among the expanded *OBP* genes, a gene named *BdioOBP45* was found to be highly expressed specifically in the ovipositor, and functional verification through *in vitro* and *in vivo* experiments indicated its potential significance in the host-seeking behavior of female parasitic wasps. This work not only provides new genome sequences for Hymenoptera systematics but also supplies a basis for the further protection and utilization of parasitoid resources in pest control through the utilization of *B. dioryctriae* genomic resources.

Methods

Insect rearing and sampling for sequencing

B. dioryctriae were obtained from the Research Institute of Forest Protection, Jilin Provincial Academy of Forestry Sciences, and then maintained at the School of Life Sciences, Northeast Normal University, Changchun, China. Individuals of *B. dioryctriae* were cultivated in a climate room with regulated circumstances (25 ± 1 °C, $50 \pm 5\%$ R. H., and 16 h L: 8 h D photoperiod) with pupae of *G. melonella* as the usual host. Following eclosion, the wasps

were raised under the same conditions as previously described while being fed a 10% (V/V) sucrose solution. Large-scale rearing was conducted in the laboratory, yielding a large population. A total of 1000 wasps that emerged within 12 h were collected for subsequent experiments to ensure the uniformity of the mixed sample.

Sequencing and genome size estimation

A modified cetyltrimethylammonium bromide (CTAB) technique was used to extract DNA from 1000 adult male *B. dioryctriae*. Using an Illumina VAHTSTM Fg DNA Library Prep Kit, a paired-end (PE) library was created, and 150 bp of PE sequencing was performed on an Illumina NovaSeq 6000 platform (Illumina, CA, USA). Using the SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences), a SMRTbell library was created, and long-read sequencing was performed on the PacBio Sequel II platform. To improve the main genome assembly to the chromosomal level, Hi-C fragment libraries were generated from 300 to 700 bp inserts using adult male samples, and the samples were sequenced using the Illumina platform [72]. Using adult males, a PE RNA-seq library was created, and the PromethION48 platform was used for the sequencing process. Genome size was estimated by K-mer analysis using short reads generated from the Illumina platform with a short insert size of 350 bp and a K-mer size of 19.

Genome assembly and Hi-C scaffolding

To create the genome sequences, high-accuracy circular consensus sequencing (CCS) data were assembled using hifiasm (v 0.16) [73] (The detailed information of the software and database involved in genomic analysis is all organized in Additional file 2: Table S14). After the Hi-C sequencing data were preliminarily filtered and evaluated using Hi-C-Pro v2.10.0 [74], for anchored contigs, BWA v0.7.10-r789 [75] was used to align the clean reads to the draft genome. Next, the contigs were clustered using the hierarchical clustering method in Lachesis, and subsequently, Lachesis was subsequently applied to order and orient the clustered contigs to perform the Hi-C-associated scaffolding [76]. By matching short reads from the Illumina platform to the genome and utilizing a set of 1367 insect BUSCOs with the BUSCO v5.2.2 software [77], the completeness of the genome assembly was evaluated.

To further confirm the chromosome number of *B. dioryctriae*, karyotype analysis was conducted. A modified variant of the procedure was used to acquire chromosome preparations from the cerebral ganglia of prepupae [78]. Images were obtained using an Olympus

orthotopic fluorescence microscope (Olympus Corp., Tokyo, Japan).

Repeat annotation

Using both de novo and homology-based methods, repetitive sequences and TEs were found across the genome. The first step was to use RepeatModeler v2.0.1 [79] to customize a de novo repeat library of the genome. LTRharvest v1.5.9 [80] and LTRfinder v2.8 [81] were then used to create high-quality intact FL-LTR-RTs and non-redundant LTR libraries. The last TE sequences in the genome of *B. dioryctriae* were located and categorized using RepeatMasker v4.12 [82] to perform a homology search against the library above with the known Dfam v3.5 database [83]. Tandem Repeats Finder TRF 409 [84] and the MicroSatellite identification tool MISA v2.1 [85] were used to annotate tandem repeats.

Prediction and functional annotation of protein-coding genes

Three techniques were used to predict gene structures: homology-based (Hisat v2.1.0 [86], and Stringtie v2.1.4 [87]), transcriptome-based (Hisat v2.1.0 [86]), and de novo (Augustus v3.1.0 [88], and SNAP 2006–07–28 [89]). In the homology-based method, the genome of *B. dioryctriae* was first mapped using GeMoMa v1.7 [90] to contain protein sequences from nine sequenced Hymenoptera insect species: *Apis mellifera*, *A. gifuensis*, *C. cunea*, *E. hayati*, *Gonatopus flavifemur*, *N. vitripennis*, *P. puparum*, and *Trichogramma pretiosum*. PASA v2.4.1 [91] updated the gene models that were merged from these various approaches using EVM software v1.1.1 [92]. Genes related to transfer RNA (tRNA) were found using the tRNAscan-SE v1.3.1 [93] method, whereas genes related to ribosomal RNA (rRNA) were found using the barrnap v0.9 [94] program with default parameters. Furthermore, Infernal v1.1 [95] was used to identify microRNAs (miRNAs) and small nuclear RNAs (snRNAs) using default parameters against the Rfam v14.5 [96].

Gene functional annotation was performed based on homolog searches and the best matches to the databases of the National Center for Biotechnology Information (NCBI) Non-Redundant (NR), Evolutionary genealogy of genes: Non-supervised Orthologous Groups (EggNOG), GO, EuKaryotic Orthologous Groups (KOG), TrEMBL, Pfam, KEGG and Swiss-Prot protein databases using diamond Basic Local Alignment Search Tool Protein (BLASTP) v2.0.4.142 [97] with an *E*-value threshold of $1E^{-5}$.

Chromosomal synteny analysis

The gene sequences of *B. dioryctriae*, *C. cunea*, and *N. vitripennis* were aligned using diamond v0.9.29.130 [98] to detect comparable gene pairings (*E* value $< 1e^{-5}$,

$P < 0.05$) in order to discover collinear gene blocks among these three species. A C score greater than 0.5 was employed as a filter criterion (the C score was filtered using JCVI software), and MCScan [99] in JCVI v0.9.13 [100] with default parameters was utilized to show high-quality blocks.

Comparative genomics

Using OrthoFinder v2.4 [101], orthologous and paralogous genes were identified in *B. dioryctriae* and 13 additional insect species (Additional file 2: Table S3), and were subsequently annotated using the PANTHER v18.0 [102]. The single-copy orthologous genes' protein sequences were aligned using the MAFFT v7.205 software [103]. The bootstrap phylogenetic tree was constructed by ML using IQTree v1.6.11 [104], and the best model, "JTTDCMut+F+I+G4," was estimated by ModelFinder. Divergence times were computed using the MCMCTREE package included in PAML v4.9i [105]. The calibration times were taken from the TimeTree website (<http://www.timetree.org/>) and were based on the following six cases: *A. rosae* vs. *C. cunea* (223.7–304.0 Ma), *C. cunea* vs. *O. abietinus* (187.9–272.5 Ma), *O. biroi* vs. *A. gifuensis* (162.4–219.3 Ma), *A. gifuensis* vs. *T. pretiosum* (139.2–253.9 Ma), *E. hayati* vs. *T. pretiosum* (107.4–241.6 Ma), and *O. biroi* vs. *A. mellifera* (100.3–163.5 Ma). The graphical representation was generated using MCMCTreeR v1.1 [106].

Gene family expansion and contraction were assessed using CAFE v4.2 [107] based on the identified gene families and the generated phylogenetic tree with the projected divergence periods of those species. Gene families with notable expansion and contraction with Viterbi $P < 0.05$ were subjected to KEGG pathway enrichment and GO analysis using the R program cluster Profiler.

Identification and localization of BdioOBP genes in the genome

For OBP gene annotation, a set of reference protein sequences (Additional file 2: Table S15) was used to identify candidate BdioOBP sequences through BLASTx (*E* value $< 1e^{-5}$) in the genome of *B. dioryctriae*. Then, the candidate sequences were identified using PfamScan against the Pfam database. Chromosome genes were mapped and visualized using TBtoolsII's "Gene Location Visualization from GTF/GFF" function [108].

Heatmap construction

Using Bowtie, the sequencing reads were aligned with the unigene dataset, and expectation maximization (RSEM) was used to assess the expression levels in conjunction with RNA-seq. The expression abundance of the relevant gene transcripts was represented by the FPKM. Using

$\log_{10}(\text{FPKM} + 1)$ values [109], heatmaps of differentially expressed *BdioOBPs* were created using the OmicShare application (<https://www.omicshare.com>).

RT-qPCR analysis and validation

RT-qPCR was used to confirm the expression patterns of the highly expressed gene *BdioOBP45* in different tissues. To standardize the expression data, ribosomal protein L18 (*BdioRPL18*) was used as a reference gene. Primer 3 (<https://bioinfo.ut.ee/primer3-0.4.0/>) was used to design specific primers, which are listed in Additional file 2: Table S16. Subsequently, qPCR was carried out in accordance with the manufacturer's instructions using the LightCycler 480 II Detection System (Roche, Shanghai, China) and TransStar Tip Top Green qPCR Supermix (TransGen Biotech, Beijing, China). Each qPCR reaction was used according to the manufacturer's instructions in a 20- μL reaction volume (0.8 μL of primer (forward and reverse, each 0.2 μM), 10 μL of 2 \times TransStar Top Green qPCR SuperMix, and 9.2 μL of RT product), and the conditions were as follows: 94 $^{\circ}\text{C}$ for 30 s, followed by 45 cycles of 94 $^{\circ}\text{C}$ for 5 s, 55 $^{\circ}\text{C}$ for 15 s, and 72 $^{\circ}\text{C}$ for 10 s. Then, 95 $^{\circ}\text{C}$ for 5 s, 65 $^{\circ}\text{C}$ for 1 min, 97 $^{\circ}\text{C}$ for 10 s, and 60 $^{\circ}\text{C}$ for 15 s were used to measure the melt curve. The RT-qPCR analysis was conducted using SPSS Statistics 26.0 to process $2^{-\Delta\Delta\text{CT}}$ results. Comparisons of the quantitative expression levels of *BdioOBP45* in different tissues were performed using one-way ANOVA, followed by Tukey's post hoc analysis ($P < 0.05$). Three technical and three biological replicates were used for each qPCR.

Gas chromatography–mass spectrometry analysis

The volatile compounds of *Pinus koraiensis* cones, pest-damaged *P. koraiensis* cones, *P. koraiensis* oleoresin subjected to mechanical damage, *P. koraiensis* oleoresin subjected to pest damage, and *D. abietella* pupae were analyzed by solid-phase microextraction (SPME). The SPME was performed using an autosampler equipped with a 2-cm 50/30 μm SPME fiber (Supelco, Bellefonte, PA, USA). TRACE 1310 gas chromatograph and TSQ 8000 mass spectrometer (Thermo Fisher Scientific, USA) were used for the analysis. Standard autotunes using perfluorotributylamine (PFTBA) were performed daily [110]. The separation conditions for the DB-5MS capillary column (Agilent Technologies, Santa Clara, CA, USA) were as follows: initial column temperature of 40 $^{\circ}\text{C}$ for 2 min, followed by a 6 $^{\circ}\text{C}/\text{min}$ climb to 250 $^{\circ}\text{C}$, and a 4-min maintenance period. The carrier gas was high-purity helium (99.999%) flowing at a rate of 1 mL/min. Compounds detected in the analyzed samples were identified using the NIST 2014 mass spectral library. Each type of sample is measured with three replicate samples, and a

blank control group is set up to ensure the quality of the detection.

Heterologous expression and purification of *BdioOBP45*

The full open reading frame of the *BdioOBP45* gene was amplified by PCR using gene-specific primers (Additional file 2: Table S16) designed based on *B. dioryctriae* genome data. Putative full-length *BdioOBP45* was sub-cloned and inserted into the expression vector pET30a (Novagen, Germany) with the restriction enzymes *EcoRI* and *XhoI* (New England BioLabs, Ipswich, England). The pET30a/*BdioOBP45* plasmid was transformed into BL21 (DE3) for expression. After induction with 0.8 mM IPTG, the proteins were expressed and collected from the culture supernatant. Protein purity was confirmed by SDS-PAGE, and to avoid any alterations in protein folding, the 6 \times His tag was deleted. For additional testing, the purified proteins were dissolved in buffer B (50 mM Tris-HCl, pH 7.4).

Fluorescence competitive binding assays

Using N-phenyl-1-naphthylamine (1-NPN) as a fluorescently labeled ligand, the binding ability of *BdioOBP45* to 55 volatiles (Additional file 2: Table S13) was measured with a Fluoromax-4 fluorescence spectrometer. The excitation wavelength was set to 337 nm, and the emission spectrum was recorded between 380 and 520 nm. *BdioOBP45* (2 μM dissolved in 50 mM Tris-HCl, pH=7.4) and 1-NPN (2–20 μM dissolved in chromatographic methanol) were mixed, and then adding aliquots of a 1 mM methanol solution of ligand to final concentrations of 2–20 μM , record the strongest of each time fluorescence value. The dissociation constants of the competitors were calculated according to previous studies [111] and analyzed by GraphPad Prism v8.0. The results of three repetitions represent the mean \pm SEM.

Y-tube olfactometer assay

To examine the biological effect of 1-octen-3-ol on *B. dioryctriae*, a binary Y-tube olfactometer (1.1 cm in diameter, 6 cm in base, 6 cm in arm length, and 60 $^{\circ}$ angle between the arms) was used for this assay. One arm of the odorant source in this assay was 20 \times 50 mm filter paper that contained 20 μL of the test chemicals. The other arm of the olfactometer was 20 \times 50 mm filter paper that contained 20 μL of the paraffin oil. When a wasp penetrated an arm by at least 2/3 and remained there for 30 s, it was considered a "choice." The data were classified as "no choice" if an insect did not decide within 5 min. The choice rate was calculated as $S_r = \frac{\text{number of individuals with positive tropism}}{\text{all the tested wasps} - \text{wasps that did not make a choice}} \times 100\%$ [38]. The tube was cleaned with dehydrated alcohol after each

of the five wasp experiments and allowed to air dry, and the locations of the odor sources were switched. The differences in Y-tube decision behavior between the paraffin oil and test chemical groups were analyzed using a chi-square test.

RNA interference

As directed by the manufacturer, PCR products were transcribed in vitro using the T7 RiboMAX Express RNA interference (RNAi) system (Promega, USA) to produce the dsRNAs of BdioOBP45. To accomplish efficient silence, the newly emerging insects were injected on the first day of the adult stage, while their olfactory system was still developing. 460 ng of dsRNA were injected into the thorax using a NanoLiter 2000 injector. For the negative control, the green fluorescent protein (GFP) gene was utilized as an irrelevant gene. Using the same method as previously used to confirm the expression of olfactory genes, RT-qPCR was employed to measure the effectiveness of gene knockdown at various intervals (12 h, 24 h, 36 h, and 48 h) following interference.

Abbreviations

Hi-C	High-through chromosome conformation capture
OSN	Olfactory sensory neuron
OBP	Odorant-binding protein
CSP	Chemosensory protein
SNMP	Sensory neuron membrane protein
NPC2	Niemann-Pick type C2 protein
OR	Odorant receptor
3-HA	3-Hexenyl acetate
BUSCO	Benchmarking universal single-copy orthologs
TE	Transposable element
CDS	Coding sequence
Chr1-Chr6	Chromosome1-chromosome6
ML	Maximum likelihood
IR	Ionotropic receptor
GR	Gustatory receptor
GO	Gene ontology
KEGG	Kyoto encyclopedia of genes and genomes
CLR	C-type lectin receptor
NLR	NOD-like receptor
FPKM	Fragments per kilobase million
RT-qPCR	Real-time quantitative polymerase chain reaction
1-NPN	1-N-phenyl-naphthylamine
Kd	Dissociation constant
GC-MS	Gas chromatography-mass spectrometry
PRR	Pattern recognition receptor
PAMP	Pathogen-associated molecular pattern
DAMP	Damage-associated molecular pattern
HIPV	Herbivore-induced plant volatile
CTAB	Cetyltrimethylammonium bromide
PE	Paired-end
CCS	Circular consensus sequencing
tRNA	Transfer RNA
rRNA	Ribosomal RNA
miRNA	MicroRNA
snRNA	Small nuclear RNA
NCBI	National center for biotechnology information
NR	Non-redundant
EggNOG	Evolutionary genealogy of genes: Non-supervised Orthologous Groups

KOG	EuKaryotic Orthologous Groups
BLAST	Basic local alignment search tool
RSEM	RNA-seq by expectation maximization
SPME	Solid-phase microextraction
PFTBA	Perfluorotributylamine
RNAi	RNA interference
GFP	Green fluorescent protein

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12915-024-01998-8>.

Additional file 1: Figure S1. Genome size of *B. dioryctriae* estimated by k-mer analysis. The k-mer ($k = 19$) frequency profile was generated by GenomeScope v1.0.0, showing an estimated genome size of 486.86 Mb. Figure S2. Genome-wide Hi-C interaction maps of six pseudochromosomal linkage groups in *B. dioryctriae*. Figure S3. Expression levels of *BdioOBP45* in different tissues. (FA and MA: female and male antennae; FH and MH: female and male heads without antennae; Fab: female abdomens without ovipositors and digestive tracts; Fov: female ovipositors; Mge: male genitalia; Mab: male abdomens without genitalia and digestive tracts; T: male and female thoraxes; L: male and female legs). Columns labeled with different letters are significantly different ($p < 0.05$). Figure S4. Expression and purification of BdioOBP45. M, protein marker; 1, before induction with IPTG; 2, after induction with IPTG; 3, after purification of the original protein; 4, after the His-tag was removed from the final protein. Figure S5. Chromatographic separation was carried out on a DB-5MS column. (A): TIC of *P. koraiensis* cones; (B): TIC of Pest-damaged *P. koraiensis* cones; (C): TIC of *P. koraiensis* oleoresin after mechanical damage; (D): TIC of *P. koraiensis* oleoresin after pest damage; (E): TIC of *D. abietella* pupae.

Additional file 2: Table S1. Genome sequencing of *B. dioryctriae* using PacBio. Table S2. Characteristics of repeat sequences in *B. dioryctriae*. Table S3. Published genomes of 13 Hymenoptera species used for analysis. Table S4. Comparison of the assembled genome size and gene structure prediction among 9 hymenopteran insects. Table S5. Functional annotation of *B. dioryctriae* genome. Table S6. Chromosome synteny based on CDS pairwise alignment between *B. dioryctriae* and *C. cunea*. Table S7. Chromosome synteny based on CDS pairwise alignment between *B. dioryctriae* and *N. vitripennis*. Table S8. Gene family clusters of fourteen species genes. Table S9. Expanded gene family of *B. dioryctriae*. Table S10. Contracted gene family of *B. dioryctriae*. Table S11. Sequence information of identified OBPs in *B. dioryctriae*. Table S12. Identification of volatile compounds from *P. koraiensis* cones; Pest-damaged *P. koraiensis* cones; *P. koraiensis* oleoresin by mechanical damage; *P. koraiensis* oleoresin by pest damage and *D. abietella* pupae. Table S13. Detailed information of the Compounds for Testing against BdioOBP45. Table S14. Software modes and database sources. Table S15. The protein names and sequences of OBPs that were used in identification candidate BdioOBPs. Table S16. Specific primers of BdioOBP45.

Acknowledgements

The authors would like to thank all the reviewers who participated in the review, as well as AJE (<https://www.aje.cn/>) for providing English editing services during the preparation of this manuscript.

Authors' contributions

QC and BZR conceived and designed the experiments. XYZ, QYL, JL, YHZ, HBJ, and LWS collected all samples. XYZ and YY performed the investigation. XYZ, YY, JL, and LD analyzed the data. XYZ and QC wrote the manuscript. All authors have read and approved the final manuscript.

Funding

This research was supported by the National Natural Science Foundation (32302432), the Science and Technology Development Program of Jilin Province (20230202096NC), the Fundamental Research Funds for the Central Universities (2412021QD010, 135111010), and the Young Science and Technology Talent Support Project of Jilin Province (QT202121).

Availability of data and materials

All the data generated in this study has been available at the National Center for Biotechnology Information (NCBI), under BioProject number PRJNA1099397 [112]. The genome assembly has been deposited at GenBank under accession JAGTPF000000000 [113]. The *B. dioryctriae* transcriptome data for genome annotation are available through NCBI BioProject (PRJNA881758) [114].

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Jilin Provincial Key Laboratory of Animal Resource Conservation and Utilization, School of Life Sciences, Northeast Normal University, Changchun 130024, China. ²Key Laboratory of Vegetation Ecology, Ministry of Education, Northeast Normal University, Changchun 130024, China. ³Jilin Provincial Engineering Laboratory of Avian Ecology and Conservation Genetics, Northeast Normal University, Changchun 130024, China. ⁴Research Institute of Forest Protection, Jilin Provincial Academy of Forestry Sciences, Changchun 130033, China. ⁵Jilin Provincial International Cooperation Key Laboratory for Biological Control of Agricultural Pests, Changchun, China.

Received: 23 April 2024 Accepted: 29 August 2024

Published online: 11 September 2024

References

- Ye X, Yan Z, Yang Y, Xiao S, Chen L, Wang J, et al. A chromosome-level genome assembly of the parasitoid wasp *Pteromalus puparum*. *Mol Ecol Resour*. 2020;20(5):1384–402.
- Pan L, Gao W, Liu X, Qin D, Zhang T, Ren R, et al. Parasitoids as taxonomists: how does the parasitoid *Chouioia cunea* distinguish between a host and a non-host? *Pest Manage Sci*. 2023;79(11):4547–56.
- Ivezić A, Trudić B. Parasitoids of the genus *Trichogramma* (Hymenoptera: Trichogrammatidae), natural enemies of European corn borer *Ostrinia nubilalis* (Hübner, 1796) (Lepidoptera: Crambidae). *J Cent Eur Agric*. 2021;22(4):787–97.
- Branstetter MG, Childers AK, Cox-Foster D, Hopper KR, Kapheim KM, Toth AL, et al. Genomes of the Hymenoptera. *Curr Opin Insect Sci*. 2018;25:65–75.
- Chen L, Lang K, Bi S, Luo J, Liu F, Ye X, et al. WaspBase: a genomic resource for the interactions among parasitic wasps, insect hosts and plants. *Database*. 2018;2018:081.
- Gokhman VEJIS. Karyotypes of parasitic Hymenoptera: Diversity, evolution and taxonomic significance. *Insect Sci*. 2006;13.
- Dalla Benetta E, Antoshechkin I, Yang T, Nguyen HQM, Ferree PM, Akbari OS. Genome elimination mediated by gene expression from a selfish chromosome. *Sci Adv*. 2020;6(14):eaa29808.
- Li B, Du Z, Tian L, Zhang L, Huang Z, Wei S, et al. Chromosome-level genome assembly of the aphid parasitoid *Aphidius gifuensis* using Oxford Nanopore sequencing and Hi-C technology. *Mol Ecol Resour*. 2021;21(3):941–54.
- Wittmeyer KT, Oppenheim SJ, Hopper KR. Assemblies of the genomes of parasitic wasps using meta-assembly and scaffolding with genetic linkage. *G3 (Bethesda)*. 2022;12(1):jkab386. <https://academic.oup.com/g3journal/article/12/1/jkab386/6423991>.
- Pinto BJ, Weis JJ, Gamble T, Ode PJ, Paul R, Zaspel JM. A Chromosome-Level Genome Assembly of the Parasitoid Wasp, *Cotesia glomerata* (Hymenoptera: Braconidae). *J Hered*. 2021;112(6):558–64.
- Gauthier J, Boulain H, van Vugt J, Baudry L, Persyn E, Aury JM, et al. Chromosomal scale assembly of parasitic wasp genome reveals symbiotic virus colonization. *Commun Biol*. 2021;4(1):104.
- Liu JF, Zhao HY, Song YF, Yu YC, Yang MF. A Chromosome-Level Genome Assembly of the Parasitic Wasp *Chelonus formosanus* Sonan 1932 (Hymenoptera: Braconidae). *Genome Biol Evol*. 2022;14(1):evac006.
- Wang Z, Ma X, Zhu J, Zheng B, Yuan R, Lu Z, et al. Chromosome-level genome assembly of *Chouioia cunea* Yang, the parasitic wasp of the fall webworm. *Sci Data*. 2023;10(1):485.
- Zhong YW, Fan YY, Zuo ZQ, Shu RG, Liu YQ, Luan JB, et al. A chromosome-level genome assembly of the parasitoid wasp *Eretmocerus hayati*. *Sci Data*. 2023;10(1):585.
- Shu X, Yuan R, Zheng B, Wang Z, Ye X, Tang P, et al. Chromosome-level genome assembly of *Microplitis manilae* Ashmead, 1904 (Hymenoptera: Braconidae). *Sci Data*. 2023;10(1):266.
- Rago A, Gilbert DG, Choi JH, Sackton TB, Wang X, Kelkar YD, et al. OGS2: genome re-annotation of the jewel wasp *Nasonia vitripennis*. *BMC Genomics*. 2016;17(1):678.
- Xu H, Ye X, Yang Y, Yang Y, Sun YH, Mei Y, et al. Comparative Genomics Sheds Light on the Convergent Evolution of Miniaturized Wasps. *Mol Biol Evol*. 2021;38(12):5539–54.
- Xiao S, Ye X, Wang S, Yang Y, Fang Q, Wang F, et al. Genome assembly of the ectoparasitoid wasp *Theocolax elegans*. *Sci Data*. 2023;10(1):159.
- Mao M, Simmonds TJ, Stouthamer CM, Kehoe TM, Geib SM, Burke GR. A chromosome scale assembly of the parasitoid wasp *Venturia canescens* provides insight into the process of virus domestication. *G3 (Bethesda)*. 2023;13(10):jkad137. <https://academic.oup.com/g3journal/article/13/10/jkad137/7204861>.
- Yin C, Li M, Hu J, Lang K, Chen Q, Liu J, et al. The genomic features of parasitism, polyembryony and immune evasion in the endoparasitoid wasp *Macrocentrus cingulum*. *BMC Genomics*. 2018;19(1):1–18.
- Wong SS, Ooi PAC, Wong WLJ. Antennal and ovipositor sensilla of *Pseudoligotia yasumatsui* (Hymenoptera: Trichogrammatidae). *J Asia-Pac Entomol*. 2019;22(1):296–307.
- Huang Z-Y, Zhang Y-J, Liu J-Y, Yang Z-D, Lu W, Zheng X-LJM, et al. Ultrastructure of female antennal sensilla of an endoparasitoid wasp, *Quadrastichus mendeli* Kim and La Salle (Hymenoptera: Eulophidae: Tetrastichinae). *Microsc Microanal*. 2018;24(4):431–441.
- Blažytė-Cėreškienė L, Aleknavičius D, Apšegaitė V, Būda V. Response of Parasitic Wasp *Cotesia glomerata* L. (Hymenoptera: Braconidae) to Cabbage Plants of Two Varieties: Olfactory Spectra of Males and Females. *J Econ Entomol*. 2022;115(5):1464–1471.
- Li Y, Chen H, Liang X, Wang S, Zhu H, Yan M, et al. Identification of candidate chemosensory genes by antennal transcriptome analysis in an ectoparasitoid wasp. *J Appl Entomol*. 2022;146(3):335–51.
- Guo H, Mo BT, Li GC, Li ZL, Huang LQ, Sun YL, et al. Sex pheromone communication in an insect parasitoid, *Campoplex chloridae* Uchida. *Proc Natl Acad Sci U S A*. 2022;119(49):e2215442119.
- Shan S, Song X, Khashaveh A, Wang SN, Lu ZY, Hussain Dhilloo K, et al. A female-biased odorant receptor tuned to the lepidopteran sex pheromone in parasitoid *Microplitis mediator* guiding habitat of host insects. *J Adv Res*. 2023;43:1–12.
- Zhu X, Liu S, Xu CQ, Guo K, Xu R, Qiao HL, et al. Morphology and distribution of sensilla on the antennae and ovipositor of the Hymenopteran parasitoid *Pseudotorymus jaapiellae* Yang et Chen. *Microsc Res Tech*. 2021;84(8):1873–86.
- Zhu X, Yu Q, Gan X, Song L, Zhang K, Zuo T, et al. Transcriptome Analysis and Identification of Chemosensory Genes in *Baryscapus dioryctriae* (Hymenoptera: Eulophidae). *Insects*. 2022;13(12):1098.
- Yadav P, Borges RM. The insect ovipositor as a volatile sensor within a closed microcosm. *J Exp Biol*. 2017;220(9):1554–1557.
- Glaser N, Gallot A, Legeai F, Montagné N, Poivet E, Harry M, et al. Candidate chemosensory genes in the Stem-borer *Sesamia nonagrioides*. *Int J Biol Sci*. 2013;9(5):481–95.
- Faucheux MJ. Embryology. Multiporous sensilla on the ovipositor of *Monopis crocipitella* Clem. (Lepidoptera: Tineidae). *Int J Insect Morphol Embryol*. 1988;17(6):473–475.
- Faucheux MJ. Embryology. Morphology and distribution of sensilla on the cephalic appendages, tarsi and ovipositor of the European sunflower moth, *Hormoia eosoma nebulella* Den. & Schiff. (Lepidoptera: Pyralidae). *Int J Insect Morphol Embryol*. 1991;20(6):291–307.

33. Klinner CF, König C, Missbach C, Werckenthin A, Daly KC, Bisch-Knaden S, et al. Functional Olfactory Sensory Neurons Housed in Olfactory Sensilla on the Ovipositor of the Hawkmoth *Manduca sexta*. *Front Ecol Evol.* 2016;4.
34. Li RT, Huang LQ, Dong JF, Wang CZ. A moth odorant receptor highly expressed in the ovipositor is involved in detecting host-plant volatiles. *Elife.* 2020;9.
35. He M, Chen H, Yang X, Gao Y, Lu Y, Cheng D. Gut bacteria induce oviposition preference through ovipositor recognition in fruit fly. *Commun Biol.* 2022;5(1):973.
36. Li R, Shan S, Song X, Khashaveh A, Wang S, Yin Z, et al. Plant volatile ligands for male-biased MmedOBP14 stimulate orientation behavior of the parasitoid wasp *Microplitis mediator*. *Int J Biol Macromol.* 2022;223(Pt A):1521–9.
37. Wang Q, Liu JT, Zhang YJ, Chen JL, Li XC, Liang P, et al. Coordinative mediation of the response to alarm pheromones by three odorant binding proteins in the green peach aphid *Myzus persicae*. *Insect Biochem Mol Biol.* 2021;130: 103528.
38. Pan L, Xiang W, Sun Z, Yang Y, Han J, Wang Y, et al. CcOBP2 plays a crucial role in 3-carene olfactory response of the parasitoid wasp *Chouioia cunea*. *Insect Biochem Mol Biol.* 2020;117: 103286.
39. Rihani K, Ferveur JF, Briand L. The 40-Year Mystery of Insect Odorant-Binding Proteins. *Biomolecules.* 2021;11(4):509.
40. Marinotti O, Ngo T, Kojin BB, Chou SP, Nguyen B, Juhn J, et al. Integrated proteomic and transcriptomic analysis of the *Aedes aegypti* eggshell. *BMC Dev Biol.* 2014;14:15.
41. Costa-da-Silva AL, Kojin BB, Marinotti O, James AA, Capurro ML. Expression and accumulation of the two-domain odorant-binding protein AaegOBP45 in the ovaries of blood-fed *Aedes aegypti*. *Parasit Vectors.* 2013;6:364.
42. Ameyna DA, Chou W, Li J, Yan G, Gershon PD, James AA, et al. Proteomics reveals novel components of the *Anopheles gambiae* eggshell. *J Insect Physiol.* 2010;56(10):1414–9.
43. Zhu J, Iovinella I, Dani FR, Liu YL, Huang LQ, Liu Y, et al. Conserved chemosensory proteins in the proboscis and eyes of Lepidoptera. *Int J Biol Sci.* 2016;12(11):1394–404.
44. Xu L, Tang KY, Chen XF, Tao Y, Jiang HB, Wang JJ. Comparative transcriptomic analysis reveals female-biased olfactory genes potentially involved in plant volatile-mediated oviposition behavior of *Bactrocera dorsalis*. *BMC Genomics.* 2021;22(1):25.
45. Song L-W, Cao L-M, Li X-P, Yang Z-Q, Chen Y-Q. A new species of *Baryscapus* (Hymenoptera: Eulophidae) parasitizing pupae and larvae of two *Dioryctria* species (Lepidoptera: Pyralidae). *Ann Entomol Soc Am.* 2017;110(3):286–293.
46. Chen Q, Liu Q, Chen Y, Du L, Zhu X, Yang Y, et al. Functional Characterization of the Niemann-Pick C2 Protein BdioNPC2b in the Parasitic Wasp *Baryscapus dioryctriae* (Chalcidodea: Eulophidae). *J Agric Food Chem.* 2024;72(14):7735–48.
47. Gonthier J, Zhang Y-B, Zhang G-F, Romeis J, Collatz J. Odor learning improves efficacy of egg parasitoids as biocontrol agents against *Tuta absoluta*. *J Pest Sci.* 2023;96(1):105–17.
48. Kruidhof HM, Kostenko O, Smid HM, Vet LE. Integrating parasitoid olfactory conditioning in augmentative biological control: potential impact, possibilities, and challenges. *Front Ecol Evol.* 2019;7:84.
49. Wilson JK, Woods HA. Innate and Learned Olfactory Responses in a Wild Population of the Egg Parasitoid *Trichogramma* (Hymenoptera: Trichogrammatidae). *J Insect Sci.* 2016;16(1):110.
50. Yang Y, Ye X, Dang C, Cao Y, Hong R, Sun YH, et al. Genome of the pincer wasp *Gonatopus flavifemur* reveals unique venom evolution and a dual adaptation to parasitism and predation. *BMC Biol.* 2021;19:1–24.
51. Wang L, Wang J, Ma Y, Wan P, Liu K, Cong S, et al. Transposon insertion causes cadherin mis-splicing and confers resistance to Bt cotton in pink bollworm from China. *Sci Rep.* 2019;9(1):7479.
52. González J, Karasov TL, Messer PW, Petrov DA. Genome-wide patterns of adaptation to temperate environments associated with transposable elements in *Drosophila*. *PLoS Genet.* 2010;6(4):e1000905.
53. Woronik A, Tunström K, Perry MW, Neethiraj R, Stefanescu C, Celorio-Mancera MD, et al. A transposable element insertion is associated with an alternative life history strategy. *Nat Commun.* 2019;10(1):5757.
54. Jiang F, Yang M, Guo W, Wang X, Kang L. Large-scale transcriptome analysis of retroelements in the migratory locust, *Locusta migratoria*. *PLoS ONE.* 2012;7(7): e40532.
55. Wu N, Zhang S, Li X, Cao Y, Liu X, Wang Q, et al. Fall webworm genomes yield insights into rapid adaptation of invasive species. *Nat Ecol Evol.* 2019;3(1):105–15.
56. Zhang S, Shen S, Peng J, Zhou X, Kong X, Ren P, et al. Chromosome-level genome assembly of an important pine defoliator, *Dendrolimus punctatus* (Lepidoptera; Lasiocampidae). *Mol Ecol Resour.* 2020;20(4):1023–37.
57. Kuang JG, Han ZT, Kang MH, Ju T, Tso S, Chen XX, et al. Chromosome-level de novo genome assembly of two conifer-parasitic wasps, *Megastigmus duclouxiana* and *Megastigmus sabinae*, reveals genomic imprints of adaptation to hosts. *Mol Ecol Resour.* 2023;23(5):1142–54.
58. Yan H, Jafari S, Pask G, Zhou X, Reinberg D, Desplan C. Evolution, developmental expression and function of odorant receptors in insects. *J Exp Biol.* 2020;223(Suppl_1):jeb208215.
59. Robertson HM, Baits RL, Walden KK, Wada-Katsumata A, Schal C.. Enormous expansion of the chemosensory gene repertoire in the omnivorous German cockroach *Blattella germanica*. *J Exp Zool B Mol Dev Evol.* 2018;330(5):265–278.
60. Lemaitre B, Hoffmann JJARI. The host defense of *Drosophila melanogaster*. *Annu Rev Immunol.* 2007;25:697–743.
61. Wicherska-Pawłowska K, Wróbel T, Rybka J. Toll-like receptors (TLRs), NOD-like receptors (NLRs), and RIG-I-like receptors (RLRs) in innate immunity. TLRs, NLRs, and RLRs ligands as immunotherapeutic agents for hematopoietic diseases. *Int J Mol Sci.* 2021;22(24):13397.
62. Tang D, Kang R, Coyne CB, Zeh HJ, Lotze MT. PAMPs and DAMPs: signal 0s that spur autophagy and immunity. *Immunol Rev.* 2012;249(1):158–75.
63. Lu Y, Su F, Zhu K, Zhu M, Li Q, Hu Q, et al. Comparative genomic analysis of C-type lectin-domain genes in seven holometabolous insect species. *Insect Biochem Mol Biol.* 2020;126: 103451.
64. Xiao J-H, Yue Z, Jia L-Y, Yang X-H, Niu L-H, Wang Z, et al. Obligate mutualism within a host drives the extreme specialization of a fig wasp genome. *Genome Biol.* 2013;14:1–18.
65. Leal WS. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu Rev Entomol.* 2013;58:373–91.
66. Lu D, Li X, Liu X, Zhang Q. Identification and molecular cloning of putative odorant-binding proteins and chemosensory protein from the bethylid wasp, *Scleroderma guani* Xiao et Wu. *J Chem Ecol.* 2007;33(7):1359–75.
67. Vieira FG, Forêt S, He X, Rozas J, Field LM, Zhou JJ. Unique features of odorant-binding proteins of the parasitoid wasp *Nasonia vitripennis* revealed by genome annotation and comparative analyses. *PLoS ONE.* 2012;7(8): e43034.
68. Li YJ, Chen HC, Hong TL, Yan MW, Wang J, Shao ZM, et al. Identification of chemosensory genes by antennal transcriptome analysis and expression profiles of odorant-binding proteins in parasitoid wasp *Aulacocentrum confusum*. *Comp Biochem Physiol Part D Genomics Proteomics.* 2021;40: 100881.
69. Boncan DAT, Tsang SSK, Li C, Lee IHT, Lam H-M, Chan T-F, et al. Terpenes and Terpenoids in Plants: Interactions with Environment and Insects. *Int J Mol Sci.* 2020;21(19):7382.
70. Chen CS, Zhao C, Wu ZY, Liu GF, Yu XP, Zhang PJ. Whitefly-induced tomato volatiles mediate host habitat location of the parasitic wasp *Encarsia formosa*, and enhance its efficacy as a bio-control agent. *Pest Manag Sci.* 2021;77(2):749–57.
71. Liu CM, Matsuyama S, Kainoh Y. Synergistic effects of volatiles from host-infested plants on host-searching behavior in the parasitoid wasp *Lytopylus rufipes* (Hymenoptera: Braconidae). *J Chem Ecol.* 2019;45(8):684–92.
72. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell.* 2014;159(7):1665–80.
73. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods.* 2021;18(2):170–175.
74. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen C-J, Vert J-P, et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol Evol.* 2015;16:1–11.

75. Li H, Durbin RJB. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*. 2009;25(14):1754–1760.
76. Burton JN, Adey A, Patwardhan RP, Qiu R, Kitzman JO, Shendure J. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nat Biotechnol*. 2013;31(12):1119–1125.
77. Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, et al. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol*. 2018;35(3):543–8.
78. Imai HT, Taylor RW, Crosland MW, Crozier RH. Modes of spontaneous chromosomal mutation and karyotype evolution in ants with reference to the minimum interaction hypothesis. *Idengaku zasshi*. 1988;63(2):159–85.
79. Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, et al. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci U S A*. 2020;117(17):9451–7.
80. Ellinghaus D, Kurtz S, Willhoeft U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinf*. 2008;9:18.
81. Xu Z, Wang H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res*. 2007;35(Web Server issue):W265–268.
82. Tarailo-Graovac M, Chen N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinformatics*. 2009;Chapter 4.4:10.11–14.10.14.
83. Wheeler TJ, Clements J, Eddy SR, Hubley R, Jones TA, Jurka J, et al. Dfam: a database of repetitive DNA based on profile hidden Markov models. *Nucleic Acids Res*. 2013;41(Database issue):D70–82.
84. Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res*. 1999;27(2):573–80.
85. Beier S, Thiel T, Münch T, Scholz U, Mascher M. MISA-web: a web server for microsatellite prediction. *Bioinformatics*. 2017;33(16):2583–5.
86. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12(4):357–60.
87. Perteua M, Perteua GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol*. 2015;33(3):290–5.
88. Stanke M, Steinkamp R, Waack S, Morgenstern B. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res*. 2004;32(suppl_2):W309–12.
89. Korf I. Gene finding in novel genomes. *BMC Bioinf*. 2004;5:59.
90. Keilwagen J, Wenk M, Erickson JL, Schattat MH, Grau J, Hartung F. Using intron position conservation for homology-based gene prediction. *Nucleic Acids Res*. 2016;44(9):e89.
91. Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith RK Jr, Hannick LI, et al. Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res*. 2003;31(19):5654–66.
92. Haas BJ, Salzberg SL, Zhu W, Perteua M, Allen JE, Orvis J, et al. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the Program to Assemble Spliced Alignments. *Genome Biol*. 2008;9(1):R7.
93. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 1997;25(5):955–64.
94. Price MN, Huang KH, Alm EJ, Arkin AP. A novel method for accurate operon predictions in all sequenced prokaryotes. *Nucleic Acids Res*. 2005;33(3):880–92.
95. Nawrocki EP, Eddy SR. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*. 2013;29(22):2933–2935.
96. Griffiths-Jones S, Moxon S, Marshall M, Khanna A, Eddy SR, Bateman A. Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res*. 2005;33(Database issue):D121–124.
97. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403–10.
98. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015;12(1):59–60.
99. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res*. 2012;40(7):e49–e49.
100. Tang H, Li J, Krishnakumar V. jcv: JCVI utility libraries. In: editor. 2015. <https://github.com/zhan4429/jcvi/blob/main/README.md>.
101. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*. 2019;20(1):238.
102. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res*. 2018;47(D1):D419–26.
103. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30(4):772–80.
104. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 2017;14(6):587–9.
105. Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci*. 1997;13(5):555–6.
106. Puttick MN. MCMCTreeR: functions to prepare MCMCTree analyses and visualize posterior ages on trees. *Bioinformatics*. 2019;35(24):5321–2.
107. Han MV, Thomas GW, Lugo-Martinez J, Hahn MW. Evolution Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Mol Biol Evol*. 2013;30(8):1987–1997.
108. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol Plant*. 2020;13(8):1194–202.
109. Yang M, Sun K, Chang J. Screening differential circular RNAs expression profiles in Vulvar Lichen Sclerosus. *Biomed Eng Online*. 2022;21(1):51.
110. Prado C, Marín P, Periago JF. Application of solid-phase microextraction and gas chromatography-mass spectrometry to the determination of volatile organic compounds in end-exhaled breath samples. *J Chromatogr A*. 2003;1011(1–2):125–34.
111. Wen M, Li E, Li J, Chen Q, Zhou H, Zhang S, et al. Molecular characterization and key binding sites of sex pheromone-binding proteins from the meadow moth *loxostege sticticalis*. *J Agric Food Chem*. 2019;67(46):12685–95.
112. Bingzhong Ren. *Baryscapus dioryctriae* isolate: Bdio-NUN Genome sequencing. <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1099397>. 2024.
113. Bingzhong Ren. *Baryscapus dioryctriae* isolate Bdio-NUN, whole genome shotgun sequencing project. GenBank. (2024). <https://www.ncbi.nlm.nih.gov/nuccore/JBDMAI000000000>.
114. Bingzhong Ren. RNA-seq of *Baryscapus dioryctriae* NCBI Bioproject accession number PRJNA881758. (2023).

Publisher's Note

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