Scientific Data

https://doi.org/10.1038/s41597-026-07168-5

Article in Press

Chromosome-level genome assembly of *Rhynchium brunneum* (Fabricius, 1787) (Hymenoptera: Vespidae)

Received: 8 October 2025

Accepted: 27 March 2026

Cite this article as: Wang, J., He, S., Chen, B. *et al.* Chromosome-level genome assembly of *Rhynchium brunneum* (Fabricius, 1787) (Hymenoptera: Vespidae). *Sci Data* (2026). https://doi.org/10.1038/s41597-026-07168-5

Jing Wang, Shulin He, Bin Chen & Tingjing Li

Chromosome-level genome assembly of *Rhynchium brunneum* (Fabricius, 1787) (Hymenoptera: Vespidae)

Jing Wang[1, 2], Shulin He[1, 2], Bin Chen[1] & Tingjing Li[1]✉

The wasps of *Rhynchium* exemplify solitary vespid predators controlling Lepidopteran pests through venom-mediated paralysis, with their venom possessing significant medicinal potential. As dominant models within the species-rich subfamily Eumeninae representing 70% of Vespidae diversity, they provide critical insights into sociality evolution and biocontrol mechanisms. To boost research on Vespidae, we used PacBio long-read, short-read RNA-seq (Illumina) and Hi-C scaffolding technologies to create a high-quality chromosome-level genome assembly for *Rhynchium brunneum*, an important solitary insect. We obtained a 328.90 Mb assembly with a Scaffold N50 size of 15.98 Mb. We detected 96.2% Benchmarking Universal Single-Copy Orthologues (BUSCO) in the genome assembly, which contains 51.77% repetitive sequences and has 12,999 protein-coding genes annotated. In *R. brunneum*, we identified 173 gene expansions and 274 genes that underwent contraction or loss. The high-quality genome of *R. brunneum* provides a valuable genetic resource for future research in evolution, molecular biology, and applied studies.

Background & Summary

The *Rhynchium* are fierce predators that primarily prey on Lepidopteran pests to feed their young. After mating, solitary females construct nests, capture prey, and store them in the nests as food for their larvae, paralyzing but not killing the prey with their venom. Furthermore, their venom also harbors significant medicinal potential, positioning it as a novel source of bioactive compounds with promise for pharmacological, therapeutic, and agricultural applications[1]. Due to their large populations, rapid flight, aggressive nature, and substantial predation rates, the species of *Rhynchium* play a direct and important role in naturally controlling Lepidopteran pests in agricultural, forest, and orchard environments[2,3]. As such, they represent a significant category of beneficial insect resources in ecological systems related to agriculture and forestry.

The genus *Rhynchium* Spinola, 1806 belongs to Insecta, Hymenoptera, Vespidae. The family Vespidae is widely distributed across the six major zoogeographic regions and encompasses over 5,300 known species globally. As the most species-rich subfamily within Vespidae, Eumeninae (potter wasps) comprises approximately 204 genera and 3,800 known species[4-7]. It is characterized by diverse nesting strategies and lifestyles ranging from predominantly solitary to primitively social in a few species[8,9]. Approximately 70% of the family Vespidae is constituted by subfamily

[1]Chongqing Key Laboratory of Control and Utilization of Vector Insects, Institute of Entomology and Molecular Biology, College of Life Science, Chongqing Normal University, Chongqing 401331, China. [2] These authors contributed equally to this work: Jing Wang, Shulin He. ✉e-mail: ltjing1979@cqnu.edu.cn.

Eumeninae, featuring a wide range of body sizes and vibrant colorations, predominantly yellow, brownish-red, and black, often adorned with various colorful patterns. Given the rich diversity in morphology, adaptive growth habits, and social-behavioral spectrum within Eumeninae, they are becoming ideal models for studying genetic diversity and the evolution of sociality in insects[10]. However, the scarcity of high-quality chromosome-level genome resources has limited our understanding of their adaptive evolution. To date, ten chromosome-level Vespidae genome assemblies have been published, with eight of these assemblies relying on high-throughput chromosome conformation capture (Hi-C) technology, among which only one Eumeninae species, *Ancistrocerus nigricornis*, has been sequenced. Given the species-rich diversity within this subfamily and compared to other social vespids, the scarcity of chromosome-level genome assemblies for solitary Eumeninae species is notable.

To enhance our knowledge of the evolution and ecology of Eumeninae, we sequenced and assembled chromosome-level genome of *Rhynchium brunneum* (Fabricius, 1787) as a reference genome of the subfamily. The species *R. brunneum* of Eumeninae is widely distributed and exhibits a wide range of color variations in China and some oriental countries[11]; its adults feed on floral nectar while larvae consume lepidopteran and coleopteran larvae provisioned by parental females in nest cells, playing crucial roles in plant pollination and biological pest control within both agricultural and natural ecosystems[6,12]. We obtained a chromosome-level genome of *R. brunneum* through the combination of PacBio long reads, short-read RNA-seq (Illumina), and Hi-C data. We annotated repeats and protein-coding genes (PCGs), and conducted gene family evolution analysis. The high-quality genome of *R. brunneum* provide valuable genetic resources for exploring genome evolution of both the family Vespidae and solitary insects and contributing to the evolution and ecology research of solitary insects.

Methods

Sample collection and genome sequencing

The *R. brunneum* were collected from Nanshan Forestry Center in Jiyuan City, Henan Province, China (35.11°N, 112.60°E) in June 2023. Adult individuals were collected by an insect net, washed with phosphate-buffered saline, and then snap-frozen in liquid nitrogen.

High-molecular-weight genomic DNA was extracted from an adult male *R. brunneum* using a modified CTAB protocol (50°C pre-heated CTAB buffer, 2% β-mercaptoethanol) followed by chloroform/isoamyl alcohol (24:1) extraction and isopropanol precipitation. DNA was resuspended in EB buffer, treated with RNase A (1% v/v, 37°C, 30 min), and further purified with AMPure PB magnetic beads. DNA integrity was verified on 1% agarose gels, and concentration was quantified using a Qubit® 4.0 Fluorometer (Invitrogen, USA) with the 1x dsDNA HS kit. For PacBio Revio long-read sequencing, libraries were generated with the SMRTbell Prep Kit 3.0 (PacBio, USA). Briefly, 5 μg DNA was sheared to ~20 kb with a Megaruptor 3 system (Diagenode), end-repaired, A-tailed, and ligated to SMRTbell adapters. After exonuclease cleanup and size selection (>5 kb) with AMPure PB beads, libraries were bound to Sequel II Revio SPRQ polymerase and sequenced on a Revio platform (PacBio, USA).

For PacBio ISO-seq, total RNA was isolated using TRIzol reagent and poly-A enriched with

the Dynabeads mRNA Purification Kit (Thermo Fisher). First-strand cDNA was synthesized using the SMARTer PCR cDNA Synthesis Kit (Takara), followed by PCR amplification with PrimeSTAR GXL polymerase. Libraries were constructed with the SMRTbell Express Template Prep Kit 2.0 (PacBio), quantified on an Agilent 2100 Bioanalyzer, and sequenced on a Sequel II platform (PacBio). In parallel, short-read RNA-seq was performed using the same poly-A enriched RNA samples. Libraries were prepared with the NEBNext Ultra II RNA Library Prep Kit (Illumina) following poly-A selection, and 150-bp paired-end sequencing was performed on an Illumina NovaSeq 6000 platform.

Using fresh tissue dissected from a single male specimen of *R. brunneum* (excluding the abdomen), we performed chromosome conformation capture (Hi-C) sequencing. The tissue was vacuum-infiltrated with nuclei-isolation buffer containing 2% formaldehyde for cross-linking, quenched with glycine, snap-frozen in liquid nitrogen and ground to powder. Nuclei were released by resuspending the powder in isolation buffer. Chromatin was digested with a restriction endonuclease, end-repaired with biotin-labelled nucleotides, and ligated to cyclize interacting fragments. After de-crosslinking, DNA was sheared to 300-700 bp; biotinylated junctions were captured with streptavidin magnetic beads and converted into Illumina-compatible libraries. Library concentration and insert size were verified with Qubit 2.0 and Agilent 2100; effective concentration was quantified by qPCR. Qualified libraries were sequenced on the Illumina platform. The total data generated from the long-read sequencing was 17.63Gb, while the total data generated from the short-read sequencing was 197.70Gb.

Genome size estimation

Genome size of *R. brunneum* was estimated by flow cytometry. Fresh heads were finely chopped and incubated in nuclear isolation buffer for 10 min. The homogenate was filtered to obtain a nuclear suspension, which was stained with propidium iodide (PI) plus RNase A on ice in the dark for 0.5-1 h[13,14]. Samples were mixed with chicken erythrocyte as internal reference at matched concentrations and analyzed using a BD FACSCalibur flow cytometer with 488 nm excitation. Genome size was calculated from the fluorescence intensity ratio between sample and reference peaks using ModFit 3.0 software, revealing a genome size of 0.67 Gb for *R. brunneum* (Table S1, Fig. 1a,b,c). Subsequent genome survey analysis via k-mer frequency distribution was performed using Smudgeplot[15] on FASTQ files obtained from third-generation sequencing to infer ploidy. The analysis predicted diploidy based on the characteristic ‘smudge’ pattern, where each smudge represents a haplotype structure. Heat intensity within the smudges corresponds to the frequency of haplotype combinations across the genome. The diploid configuration (AB) exhibited the highest frequency, corroborating the diploid prediction (Fig. 1d). In Hymenoptera, males typically develop as haploids from unfertilized eggs, while females arise as diploids from fertilized eggs. However, in this study, the *R. brunneum* male was found to be diploid. Two biological mechanisms can explain this observation. Firstly, diploid males are known to occur in several hymenopteran lineages, particularly under inbreeding conditions. In the subfamily Eumeninae to which *R. brunneum* belongs, diploid males have been documented in *Euodynerus foraminatus* and *Ancistrocerus antilope*. In *E. foraminatus*, sibling mating is common, with approximately 40% of

females mating with their brothers before dispersal[16]. In *A. antilope*, 90% of matings occur between siblings, and 25% of field-collected males are diploid, consistent with high inbreeding levels[17]. Thus, diploid males can occur across species and conditions. Secondly, tissue-specific endoreduplication is widespread in Hymenoptera. Even in genetically haploid males, key somatic tissues such as flight muscles can restore diploid DNA content via endoreduplication, a phenomenon observed across nearly all hymenopteran families except the most basal Xyelidae[18]. Therefore, even if the individual were haploid, DNA extracted from whole body would contain diploid nuclei from such tissues, yielding a heterozygous pattern in Smudgeplot. Both mechanisms are biologically plausible, suggesting that our *R. brunneum* sample may indeed represent a diploid male or contain diploid somatic tissues.

Figure 1 goes here.

Genome assembly

The initial assembly was generated with PacBio Hifi sequencing data by using Hifiasm (v0.15.1)[19] with default parameters. The potential bacterial and human contaminant sequences were removed by using the NCBI Foreign Contamination Screen (FCS)[20]. Haplotypic duplicates were then purged using purge_dups[21], yielding the final *R. brunneum* genome assembly of 97 contigs with 328.90 Mb with a contig N50 of 9.96 Mb. The clean Hi-C reads were aligned to the draft genome assembly using BWA (0.7.10)[22] with default parameters. We anchored the contigs onto chromosome level scaffolds using YaHS[23]. The final assembly was manually corrected in Juicebox (v1.11.08)[24]. The use of a HiC approach, which enabled us to organize the scaffolds at the chromosome level, was successful as it resulted in a final genome of 328.90 Mb composed of 17 scaffolds and an N50 of 15.98 Mb (Figs. 1e, 2, Table 1).

We evaluated the genome assembly completeness using BUSCO (v4.1.4)[25] with the hymenoptera_odb10 orthologue data set. The analysis identified 96.3% (single-copied genes: 96.2%, duplicated genes: 0.1%), 0.7%, and 3.0% of the 5991 predicted genes in this genome as complete, fragmented, and missing sequences, respectively. These results suggested the assembled genome is highly complete.

According to published genome sequencing projects, more than 80% of sequenced Hymenoptera species have genomes ranging from 180 to 340 Mb[26]. Thus, the assembled genome size of *R. brunneum* of 328.90 Mb falls well within the typical range for the order. However, our flow cytometry based estimate of 0.67 Gb is larger than the assembly. Such discrepancies between flow cytometry and sequencing based assemblies are commonly observed across insects[27,28]. For instance, in the coleopteran species *Bembidion transversale*, flow cytometry estimated a genome size of 2,118.1 Mb, whereas the assembled genome size was only 661.39 Mb[27].

In our case, the high BUSCO completeness of 96.3% indicates that the gene set is essentially complete, suggesting the disparity likely arises from technical and biological factors rather than missing genes. Firstly, highly repetitive genomic regions such as centromeres and telomeres remain challenging to assemble accurately even with long-read technologies and may be underrepresented in assemblies. Secondly, somatic polyploidy in head tissues may also lead to an

overestimate by flow cytometry. Studies have shown that in *Drosophila*, polyploid cells accumulate in the adult brain, particularly in the optic lobes where over 20% of cells can exceed 4C DNA content[29]. Polyploidy is also widespread in other insects[30]. For example, in the honey bee *Apis mellifera*, the malpighian tubules are highly polyploid secretory cells, and the brain, stinger, leg, thoracic muscle, and flight muscle also generate polyploid cells[31]. It is therefore likely that *R. brunneum* heads contain such polyploid somatic cells, and flow cytometry would indiscriminately measure their DNA, leading to a flow cytometry estimate that is nearly double the assembled genome size. Together, these factors help explain why the flow cytometry estimate exceeds the assembly size.

Figure 2 goes here.

Elements		Current Version
Genome assembly	Assembly size (Mb)	328.90
	Number of scaffolds	147
	Longest scaffold (Mb)	22.36
	N50 scaffold length (Mb)	15.98
	GC (%)	36.61
	BUSCO completeness (%)	96.30
	Chromosome-level scaffolds	17
	Anchoring rate (%)	81.40
Gene annotation	Protein-coding genes	12,999
	BUSCO completeness (%)	96.40
	Transcripts	15,215
	Exon	83,385
	Intron	69,751

Table 1.

Genome assembly and annotation statistics of *R. brunneum*.

Genome annotation

The repeat sequences of transposable elements (TE) were annotated by using a combination of homology-based and *de novo* approaches. To predict repeats *ab initio*, we constructed a *de novo* repeat library of the genome by using RepeatModeler (v2.0.2)[32]. We used RepeatModeler to identify non-long terminal repeat (LTR) retrotransposons and the remaining unclassified TEs. RepeatMasker (v4.1.2, http://www.repeatmasker.org) was then used to search for known and novel TEs by mapping sequences against the *de novo* TE library and Repbase library (www.girinst.org). Long terminal repeat retrotransposons were searched *de novo* by using LTR_FINDER_parallel[33], LTR_HARVEST_parallel[34], and LTR_retriever (v2.8)[35]. The results of all the above analyses were integrated to generate a high quality nonredundant TE library for whole-genome TE annotation. In total, 170.29 Mb of repetitive sequences were annotated, occupying 51.77% of the entire

genome. Among these, DNA transposons accounted for 2.41%, retroelements for 44.40%, LTR elements for 43.96%, LINEs for 0.44%, unclassified elements for 4.92%, and Rolling-circles and simple repeats for very little (Table 2, Table S2).

Repeat type	Length occupied (bp)	Proportion in Genome
Retroelements	146,042,958	44.40%
DNA transposon	7,921,154	2.41%
LINE	1,456,255	0.44%
LTR	144,586,703	43.96%
Rolling-circles	109,824	0.03%
Simple repeat	45,093	0.01%
Unknown	16,168,802	4.92%
Total	170,287,831	51.77%

Table 2. Statistics of repetitive elements in the genome *R. brunneum*.

We annotated the high-quality protein-coding genes used a combination of homology-based, *de novo*, and transcriptome-based approaches. Prior to genome annotation, the repeat annotation results were used to soft mask the genomes of *R. brunneum*. We used the BRAKER3[36] pipeline that built on GeneMark-ETP[37] and AUGUSTUS[38] and further improved accuracy using the TSEBRA[39] combiner, which integrates extrinsic evidence from long-read RNA-seq (PacBio ccs), short-read RNA-seq (Illumina) and a large database of protein sequences into a single prediction. The protein sequences of Polistinae and Vespinae were downloaded from the NCBI database as references for homology-based prediction. Transcript sequences were assembled with StringTie (v2.1.6)[40] from the short RNA-seq reads aligned to the genome by HISAT2 (v2.2.1)[41]. The GeneMark-ETP module trained its model and generated predictions using assembled transcripts and homologous protein databases as input data, followed by the AUGUSTUS (v3.3.3), which incorporated the predictions from the previous step to train its parameters and performed iterative refinement of gene structures. Then the transcripts was mapped from the assembled subread libraries of long-read RNA-seq to the genome sequence and used with GeneMarkS-T[42] to identify protein-coding regions. The GeneMarkS-T coordinates and the long-read transcripts were used to create a gene set in GTF format. In the final step, the long-read version of TSEBRA was used to combine three predicted gene sets using all extrinsic evidence. A total of 12,999 protein-coding genes and 15,215 transcripts of *R. brunneum* have been annotated by combining transcriptomic data, nine known protein sets of both Polistinae and Vespinae, and *ab initio* predictions (Table 1).

InterProScan 5.52-86.0[43] was used to analysis protein domains with the available different databases in InterProScan. Both Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were annotated using eggNOGmapper (v2.0.1)[44] against the eggNOG database (v5.0.1)[45]. Gene functional annotation results derived from multiple databases revealed that 12,157 genes (93.52% of the total predicted gene set) were successfully annotated to functional databases.

Figure 3 goes here.

Phylogeny and gene family evolution

We downloaded the genomes of ten hymenopteran species from NCBI for phylogenetic analyses, including six species of Vespidae: *Ancistrocerus nigricornis* (Curtis, 1826), *Odynerus spinipes* (Linnaeus, 1758), *Polistes fuscatus* (Fabricius, 1793), *Vespa crabro* Linnaeus, 1758, *Vespa mandarinia* Smith, 1852, and *Vespula vulgaris* (Linnaeus, 1758), and *Apis mellifera* Linnaeus, 1758, *Nasonia vitripennis* (Walker, 1836), *Neodiprion lecontei* (Fitch, 1858), and *Solenopsis invicta* Buren, 1972 (Table 3; Fig. 3). We used OrthoFinder (v2.5.4)[46] to infer orthogroups with sequence alignment with the Diamond ultra-sensitive mode. A total of 121,833 genes (96.1% of the total gene set) were clustered into 11,767 orthogroups, with over 80% of genes successfully assigned to orthogroups and each orthogroup containing an average of 11 genes. There were 1,001 species-specific orthogroups, which included genes from only a single species, possibly due to species-specific gene duplication. In *R. brunneum*, 1,215 species-specific genes were found (Fig. 4c).

Figure 4 goes here.

Species	Family	Subfamily	Source
A. nigricornis	Vespidae	Eumeninae	Ensembl (GCA_916049575.1)
A. mellifera	Apidae	Apinae	NCBI (GCF_003254395.2)
N. vitripennis	Pteromalidae	Pteromalinae	NCBI (GCF_009193385.2)
N. lecontei	Diprionidae	Diprioninae	NCBI (GCF_021901455.1)
O. spinipes	Vespidae	Eumeninae	NCBI (GCA_032403825.1)
P. fuscatus	Vespidae	Polistinae	NCBI (GCF_010416935.1)
R. brunneum	Vespidae	Eumeninae	This study
S. invicta	Formicidae	Myrmicinae	NCBI (GCF_016802725.1)
V. crabro	Vespidae	Vespinae	NCBI (GCF_910589235.1)
V. mandarinia	Vespidae	Vespinae	NCBI (GCF_014083535.2)
V. vulgaris	Vespidae	Vespinae	NCBI (GCF_905475345.1)

Table 3. Species taxonomic information and accession code of all samples used in this study.

To reveal the phylogenetic relationships between *R. brunneum* and 10 other hymenopteran species, we constructed a phylogenetic tree based on single-copy orthologues, with *N. lecontei* from the suborder Symphyta designated as the outgroup relative to the apocritan taxa. Universal single-copy orthologs (USCOs) were extracted with BUSCO against Hymenoptera reference gene sets (n = 5991). The USCO amino acid and nucleotide sequences were then utilized for subsequent analyses. The USCO amino acid and nucleotide sequences of each locus were aligned using the L-INS-I strategy in MAFFT[47]. Subsequently, they were trimmed using trimAl (v1.4.1)[48] to eliminate gaps and ambiguous sites. The trimmed alignments were concatenated by FASConCAT-g (v1.05)[49], generating matrix USCO100 with 100% completeness. Phylogenetic analyses were performed

based on USCO amino acid matrix of 100% completeness (USCO100_faa). The phylogenetic tree was constructed by IQ-TREE (v2.0.5)[50]. We used the MODELFINDER[51] to select the most appropriate substitution model by employing the relaxed hierarchical clustering algorithm '-rcluster 10' in IQ-TREE. The best protein substitution model for the USCO matrices was 'LG + I + G'. The support for the resulting maximum likelihood (ML) tree was evaluated using UFBoot2[52] and SH-aLRT[53], with 1000 replicates for each. As expected, the ML phylogenetic analysis showed that the three Eumeninae species (*R. brunneum*, *A. nigricornis*, and *O. spinipes*) cluster together, with three species from Vespinae and one from Polistinae (*V. vulgaris*, *V. crabro*, *V. mandarinia* and *P. fuscatus*) forming a sister group, and *R. brunneum* has a closer relationship to *A. nigricornis* than to *O. spinipes* within the subfamily Eumeninae (Fig. 4a).

We used MCMCTree in PAML v4.9j[54] to estimate the divergence times of the species by performing approximate likelihood calculations to estimate the divergence times of the species. The results were calibrated using two standard divergence time points. Two points were obtained from the TimeTree database (http://timetree.org/): (a) *P. fuscatus-V. crabro*, 63.3-76.6 million years ago (Mya) and (b) *N. vitripennis-A. mellifera*, 162.4-219.3 Mya. To reduce the computational burden, approximate likelihood calculation and ML estimation of branch lengths was performed by using 'usedata = 2 and 3'. We used the independent rates clock model (clock = 2) and the GTR substitution model (model = 7) to calculate the Hessian matrices. We set the parameters of the mean substitution rate and the rate drift as 'rgene_gamma = 2 20 1' and 'sigma2_gamma = 1 10 1'. The estimation was run for 1 billion MCMC generations sampled every 1,000,000 generations, with the first 0.2 million as burnin[55]. Estimated divergence times of *R. brunneum* and other species suggested that *R. brunneum* diverged from the common ancestor of *A. nigricornis* in the subfamily Eumeninae approximately 36.37 Mya. *R. brunneum* and *A. nigricornis* are reciprocally monophyletic sister taxa within Eumeninae. Their divergence from *O. spinipes* occurred at approximately 62.11 Mya (Fig. 4a).

We used CAFE (v4.2.1)[56] to analyze gene family expansion and contraction during the evolution of *R. brunneum* and related species. We identified 173 expanded and 274 contracted gene families in *R. brunneum* (Fig. 4b). We performed functional enrichment analysis (GO and KEGG) on protein-coding genes (PCGs) from significantly expanded families by using ClusterProfiler (v3.10.1)[57] with default parameters. GO and KEGG enrichment analyses revealed that the expanded gene families of *R. brunneum* are enriched in olfactory receptor activity (GO: 0004984), odorant binding (GO: 0005549), dendrite membrane (GO: 0032590), neuron projection membrane (GO: 0032589), and detection of chemical stimuli involved in sensory perception of smell (GO: 0050911). These findings suggest a potential evolutionary link between gene family expansion and the chemosensory system in *R. brunneum* (Tables. S3, S4, Fig. 5a,b). However, as these inferences are based on comparative genomic data, further experimental investigations will be necessary to validate the functional implications of these expansions. Additionally, we acknowledge that the current taxonomic sampling may influence these results, and that including more representative species in the future could refine the inferred patterns of gene family evolution.

Figure 5 goes here.

Chromosome synteny

To reveal the syntenic relationships between *A. mellifera*, *V. crabro* and *R. brunneum*, MCScanX[58] was used to analyze the protein-coding genes from the genome homology alignments. *R. brunneum* showed a low level of synteny with the more distantly related hymenopteran *A. mellifera* (Fig. 5c).

Data Records

The raw sequencing data of *Rhynchium brunneum* has been deposited at the National Center for Biotechnology Information (NCBI). The PacBio, Illumina, Hi-C, and transcriptome data can be found under identification numbers SRR35603922-SRR35603931[59-68]. The assembled genome has been deposited in the NCBI assembly with the accession number GCA_055773155.1[69]. The genome annotation information has been deposited in the Figshare database[70].

Technical Validation

Two methods were used to evaluate the quality of the genome assembly. Firstly, the clean reads acquired from Illumina sequencing were aligned against the genome assembly using BWA. The results revealed that 96.24% of Illumina reads were aligned to the genome assembly. Secondly, we assessed assembly completeness using BUSCO with the insecta_odb10 ortholog data set (n = 1,367). The final genome assembly showed a BUSCO completeness of 99.0%, consisting of 1,347 (98.5%) single-copy BUSCOs, 7 (0.5%) duplicated BUSCOs, 4 (0.3%) fragmented BUSCOs, and 9 (0.7%) missing BUSCOs. These evaluations collectively reflect the high quality of the genome assembly produced in this study.

Data Availability

The raw sequencing data of *Rhynchium brunneum* has been deposited at the National Center for Biotechnology Information (NCBI). The PacBio, Illumina, Hi-C, and transcriptome data are available under accession numbers SRR35603922-SRR35603931[59-68]. The assembled genome has been deposited in the NCBI assembly with the accession number GCA_055773155.1[69]. The genome annotation information has been deposited in the Figshare database[70].

Code availability

No custom scripts or code were generated in this study. All data analyses were performed according to the manual and protocols of the published bioinformatic tools.

References

1. Lee, S. H., Baek, J. H. & Yoon, K. A. Differential Properties of Venom Peptides and Proteins in Solitary vs. Social Hunting Wasps. *Toxins* **8**, 32, https://doi.org/10.3390/toxins8020032 (2016).
2. Klein, A., Steffan-Dewenter, I. & Tscharntke, T. Foraging trip duration and density of megachilid bees, eumenid wasps and pompilid wasps in tropical agroforestry systems. *Journal of Animal Ecology* **73**, 517-525, https://doi.org/10.1111/j.0021-8790.2004.00826.x (2004).

3. Dang, H. & Nguyen, L. Nesting biology of the potter wasp *Rhynchium brunneum brunneum* (Fabricius, 1793) (Hymenoptera: Vespidae: Eumeninae) in North Vietnam. *Journal of Asia-Pacific Entomology* **22**, https://doi.org/10.1016/j.aspen.2019.02.003 (2019).
4. Carpenter, J. M. The phylogenetic relationships and natural classification of the Vespoidea (Hymenoptera). *Systematic Entomology* **7**, 11-38, https://doi.org/10.1111/j.1365-3113.1982.tb00124.x (1982).
5. Hermes, M. G., Melo, G. A. R. & Carpenter, J. M. The higher-level phylogenetic relationships of the Eumeninae (Insecta, Hymenoptera, Vespidae), with emphasis on Eumenes sensu lato. *Cladistics* **30**, 453-484, https://doi.org/10.1111/cla.12059 (2014).
6. Li, T., Barthelemy, C. & Carpenter, J. The Eumeninae (Hymenoptera, Vespidae) of Hong Kong (China), with description of two new species, two new synonymies and a key to the known taxa. *Journal of Hymenoptera Research* **72**, 127-176, https://doi.org/10.3897/jhr.72.37691 (2019).
7. Brozoski, F., de Lima, V. A., Ferrari, R. R. & Buschini, M. L. T. Nesting Biology of the Potter Wasp *Ancistrocerus flavomarginatus* (Hymenoptera, Vespidae, Eumeninae) Revealed by Trap-Nest Experiments in Southern Brazil. *Neotropical Entomology* **52**, 11-23, https://doi.org/10.1007/s13744-022-01004-2 (2023).
8. West-Eberhard, M. Behavior of the primitively social wasp Montezumia cortesioides Willink (Vespidae Eumeninae) and the origins of vespid sociality. *Ethology Ecology & Evolution* **17**, 201-215, https://doi.org/10.1080/08927014.2005.9522592 (2005).
9. Hermes, M., Somavilla, A. & Garcete Barrett, B. R. On the nesting biology of Pirhosigma Giordani Soika (Hymenoptera, Vespidae, Eumeninae), with special reference to the use of vegetable matter. *Revista Brasileira de Entomologia* **57**, 433-436, https://doi.org/10.1590/S0085-56262013005000044 (2013).
10. Kelstrup, H. C., West-Eberhard, M. J., Nascimento, F., Riddiford, L. & Hartfelder, K. Behavior, ovarian status, and juvenile hormone titer in the emblematic social wasp *Zethus miniatus* (Vespidae, Eumeninae). *Behavioral Ecology and Sociobiology* **77**, https://doi.org/10.1007/s00265-023-03334-6 (2023).
11. Peng, Y.-L., He, S.-L., Chen, B. & Li, T.-J. An Integrative Phylogenetic Analysis of the Genus *Rhynchium* Spinola (Hymenoptera: Vespidae: Eumeninae) from China Based on Morphology, Genomic Data and Geographical Distribution. *Insects* **16**, https://doi.org/10.3390/insects16020217 (2025).
12. Fateryga, A. & Amolin, A. Nesting and biology of *Jucancistrocerus caspicus* (Hymenoptera, Vespidae, Eumeninae). *Entomological Review* **94**, 73-78, https://doi.org/10.1134/S0013873814010084 (2014).
13. Dolezel, J. & Bartos, J. Plant DNA flow cytometry and estimation of nuclear genome size. *Annals of Botany* **95**, 99-110, https://doi.org/10.1093/aob/mci005 (2005).
14. Dolezel, J., Greilhuber, J. & Suda, J. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* **2**, 2233-2244, https://doi.org/10.1038/nprot.2007.310 (2007).
15. Ranallo-Benavidez, T. R., Jaron, K. S. & Schatz, M. C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nature Communications* **11**, 1432, https://doi.org/10.1038/s41467-020-14998-3 (2020).
16. Cowan, D. P. Sibling matings in a hunting wasp: adaptive inbreeding? *Science* **205**, 1403-1405, https://doi.org/10.1126/science.205.4413.1403 (1979).
17. Chapman, T. & C Stewart, S. Extremely high levels of inbreeding in a natural population of the free-living wasp *Ancistrocerus antilope* (Hymenoptera: Vespidae: Eumeninae). *Heredity* **76**, https://doi.org/10.1038/hdy.1996.8 (1996).

18. Aron, S., de Menten, L., Van Bockstaele, D. R., Blank, S. M. & Roisin, Y. When Hymenopteran Males Reinvented Diploidy. *Current Biology* **15**, 824-827, https://doi.org/10.1016/j.cub.2005.03.017 (2005).
19. 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, https://doi.org/10.1038/s41592-020-01056-5 (2021).
20. Astashyn, A. *et al.* Rapid and sensitive detection of genome contamination at scale with FCS-GX. *Genome Biology* **25**, 60, https://doi.org/10.1186/s13059-024-03198-7 (2024).
21. Guan, D. *et al.* Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics* **36**, 2896-2898, https://doi.org/10.1093/bioinformatics/btaa025 (2020).
22. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760, https://doi.org/10.1093/bioinformatics/btp324 (2009).
23. Zhou, C., McCarthy, S. A. & Durbin, R. YaHS: yet another Hi-C scaffolding tool. *Bioinformatics* **39**, https://doi.org/10.1093/bioinformatics/btac808 (2023).
24. Dudchenko, O. *et al. De novo* assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. *Science* **356**, 92-95, https://doi.org/10.1126/science.aal3327 (2017).
25. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210-3212, https://doi.org/10.1093/bioinformatics/btv351 (2015).
26. Mei, Y. et al. InsectBase 2.0: a comprehensive gene resource for insects. *Nucleic Acids Research* **50**, D1040-d1045, https://doi.org/10.1093/nar/gkab1090 (2022).
27. Pflug, J. M., Holmes, V. R., Burrus, C., Johnston, J. S. & Maddison, D. R. Measuring Genome Sizes Using Read-Depth, k-mers, and Flow Cytometry: Methodological Comparisons in Beetles (Coleoptera). *G3: Genes, Genomes, Genetics* **10**, 3047-3060, https://doi.org/10.1534/g3.120.401028 (2020).
28. He, K., Lin, K., Wang, G. & Li, F. Genome Sizes of Nine Insect Species Determined by Flow Cytometry and k-mer Analysis. *Frontiers In Physiology* **7**, 569, https://doi.org/10.3389/fphys.2016.00569 (2016).
29. Nandakumar, S., Grushko, O. & Buttitta, L. A. Polyploidy in the adult *Drosophila* brain. *Elife* **9**, https://doi.org/10.7554/eLife.54385 (2020).
30. Ren, D., Song, J., Ni, M., Kang, L. & Guo, W. Regulatory Mechanisms of Cell Polyploidy in Insects. *Frontiers in Cell and Developmental Biology* **8**, https://doi.org/10.3389/fcell.2020.00361 (2020).
31. Rangel, J., Strauss, K., Seedorf, K., Hjelmen, C. E. & Johnston, J. S. Endopolyploidy changes with age-related polyethism in the honey bee, *Apis mellifera*. *PLoS One* **10**, e0122208, https://doi.org/10.1371/journal.pone.0122208 (2015).
32. Flynn, J. M. *et al.* RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences* **117**, 9451-9457, https://doi.org/10.1073/pnas.1921046117 (2020).
33. Ou, S. & Jiang, N. LTR_FINDER_parallel: parallelization of LTR_FINDER enabling rapid identification of long terminal repeat retrotransposons. *Mobile DNA* **10**, 48, https://doi.org/10.1186/s13100-019-0193-0 (2019).
34. Ellinghaus, D., Kurtz, S. & Willhoeft, U. LTRharvest, an efficient and flexible software for *de novo* detection of LTR retrotransposons. *BMC Bioinformatics* **9**, 18, https://doi.org/10.1186/1471-2105-9-18 (2008).

35. Ou, S. & Jiang, N. LTR_retriever: A Highly Accurate and Sensitive Program for Identification of Long Terminal Repeat Retrotransposons. *Plant Physiology* **176**, 1410-1422, https://doi.org/10.1104/pp.17.01310 (2018).
36. Gabriel, L. *et al.* BRAKER3: Fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA. *Genome Research* **34**, 769-777, https://doi.org/10.1101/gr.278090.123 (2024).
37. Brůna, T., Lomsadze, A. & Borodovsky, M. GeneMark-ETP significantly improves the accuracy of automatic annotation of large eukaryotic genomes. *Genome Research* **34**, 757-768, https://doi.org/10.1101/gr.278373.123 (2024).
38. Stanke, M. *et al.* AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Research* **34**, W435-439, https://doi.org/10.1093/nar/gkl200 (2006).
39. Gabriel, L., Hoff, K. J., Brůna, T., Borodovsky, M. & Stanke, M. TSEBRA: transcript selector for BRAKER. *BMC Bioinformatics* **22**, 566, https://doi.org/10.1186/s12859-021-04482-0 (2021).
40. Kovaka, S. *et al.* Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biology* **20**, 278, https://doi.org/10.1186/s13059-019-1910-1 (2019).
41. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* **37**, 907-915, https://doi.org/10.1038/s41587-019-0201-4 (2019).
42. Tang, S., Lomsadze, A. & Borodovsky, M. Identification of protein coding regions in RNA transcripts. *Nucleic Acids Research* **43**, e78, https://doi.org/10.1093/nar/gkv227 (2015).
43. Jones, P. *et al.* InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236-1240, https://doi.org/10.1093/bioinformatics/btu031 (2014).
44. Huerta-Cepas, J. *et al.* Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Molecular Biology and Evolution* **34**, 2115-2122, https://doi.org/10.1093/molbev/msx148 (2017).
45. Huerta-Cepas, J. *et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research* **47**, D309-d314, https://doi.org/10.1093/nar/gky1085 (2019).
46. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology* **20**, 238, https://doi.org/10.1186/s13059-019-1832-y (2019).
47. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**, 772-780, https://doi.org/10.1093/molbev/mst010 (2013).
48. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973, https://doi.org/10.1093/bioinformatics/btp348 (2009).
49. Kück, P. & Longo, G. C. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Frontiers in Zoology* **11**, 81, https://doi.org/10.1186/s12983-014-0081-x (2014).
50. Minh, B. Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* **37**, 1530-1534, https://doi.org/10.1093/molbev/msaa015 (2020).

51. 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, https://doi.org/10.1038/nmeth.4285 (2017).
52. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* **35**, 518-522, https://doi.org/10.1093/molbev/msx281 (2018).
53. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59**, 307-321, https://doi.org/10.1093/sysbio/syq010 (2010).
54. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**, 1586-1591, https://doi.org/10.1093/molbev/msm088 (2007).
55. dos Reis, M. & Yang, Z. Approximate likelihood calculation on a phylogeny for Bayesian estimation of divergence times. *Molecular Biology and Evolution* **28**, 2161-2172, https://doi.org/10.1093/molbev/msr045 (2011).
56. Han, M. V., Thomas, G. W., Lugo-Martinez, J. & Hahn, M. W. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Molecular Biology and Evolution* **30**, 1987-1997, https://doi.org/10.1093/molbev/mst100 (2013).
57. Yu, G., Wang, L., Han, Y. & He, Q. Clusterprofiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology* **16**, 284-287, https://doi.org/10.1089/omi.2011.0118 (2012).
58. Wang, Y. *et al.* MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research* **40**, e49, https://doi.org/10.1093/nar/gkr1293 (2012).
59. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR35603922 (2025).
60. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR35603923 (2025).
61. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR35603924 (2025).
62. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR35603925 (2025).
63. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR35603926 (2025).
64. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR35603927 (2025).
65. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR35603928 (2025).
66. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR35603929 (2025).
67. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR35603930 (2025).
68. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR35603931 (2025).
69. Wang, J., He, S.L., Chen, B. & Li, T. J. Genbank https://identifiers.org/insdc.gca:GCA_055773155.1 (2026).
70. Wang, J. Chromosome-level genome assembly of *Rhynchium brunneum* (Fabricius, 1787) (Hymenoptera: Vespidae). *figshare* https://doi.org/10.6084/m9.figshare.30227221 (2025).

Acknowledgements

We are grateful to Chun-Lin He (Henan University of Science and Technology, Luoyang, China) for providing us with some specimens for this research. This study was funded by the Science & Technology Fundamental Resources Investigation Program (No. 2022FY202100) and the National Natural Science Foundation of China (No. 31772490, 31372247, 31000976).

Author contributions

T.-J.L. and B.C. contributed to the research design. J.W. and S.-L.H. analyzed the data. J.W., S.-L.H. and T.-J.L. wrote the draft manuscript and revised the manuscript. All co-authors contributed to this manuscript and approved it.

Competing interests

The authors declare no competing interests.

Figure legends

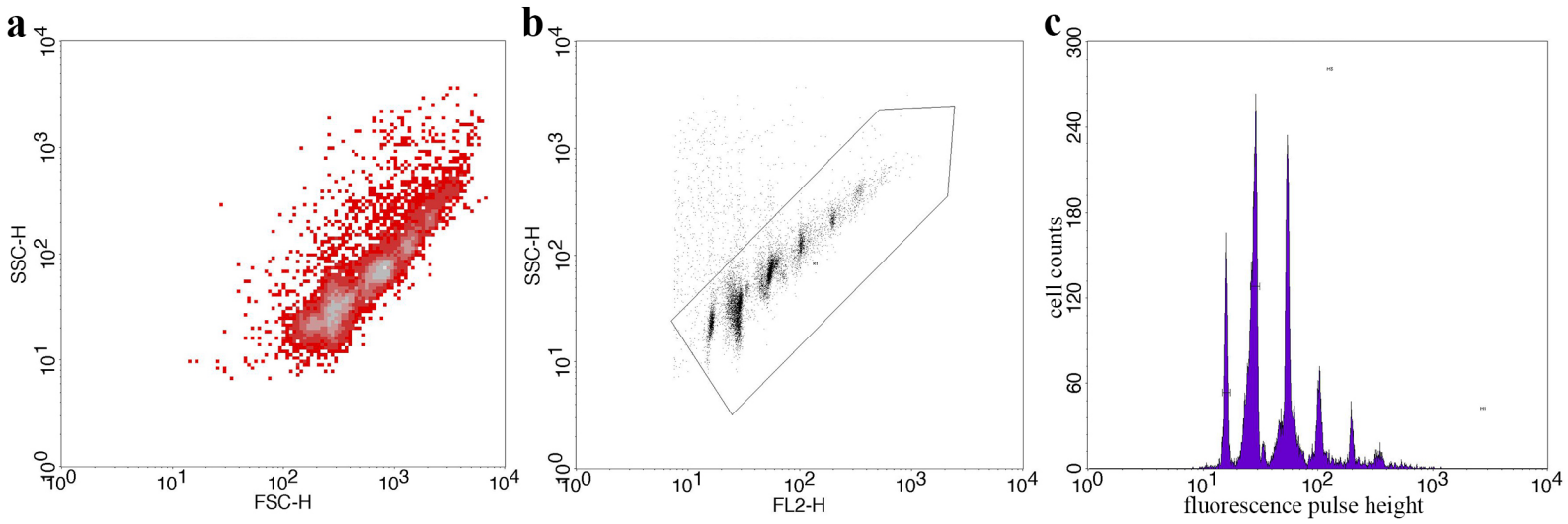
Fig. 1 (**a**) 2D density plot of *R. brunneum*. The x-axis (FSC) separates particles by size, and the y-axis (SSC) separates them by granularity and complexity. Particles with the same size and density cluster together, and areas with more clustered particles appear lighter in color; (**b**) 2D dot plot of *R. brunneum*. The y-axis uses SSC to separate particles by granularity and complexity. The x-axis (FL2) separates particles by fluorescence intensity. Fluorescence intensity correlates positively with sample DNA content; (**c**) Histogram of *R. brunneum*. The first peak is the *R. brunneum* sample peak; the others are internal reference peaks; (**d**) Smudgeplot heatmap of *R. brunneum*. The x-axis represents relative coverage (CovB / (CovA + CovB)), the y-axis represents total coverage (CovA + CovB), and the color indicates the frequency of k-mer pairs; (**e**) Genome-wide interaction heatmap of *R. brunneum*, with chromosome-level scaffolds and contig framed in blue and green, respectively.

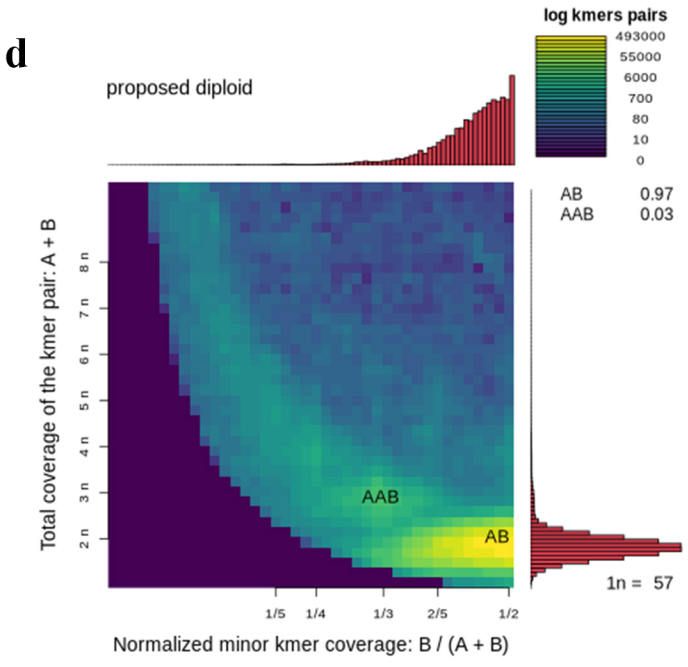
Fig. 2 Genome landscape of *R. brunneum*. From outer to inner circles: I, seventeen chromosome-level scaffolds at the Mb scale; II, gene density across the genome; III, GC contents across the genome; IV, GC-skew. The photo in the centre is of *R. brunneum.*

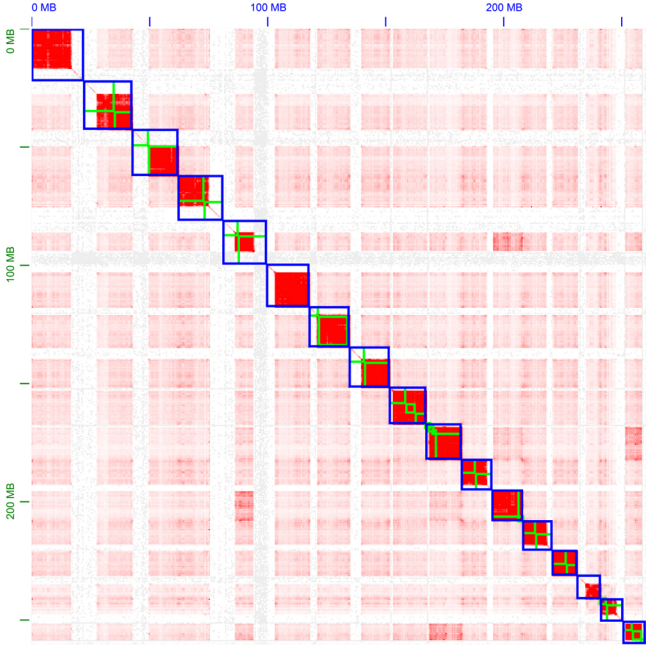
Fig. 3 BUSCO completeness: complete (C) and single-copy (S, Light blue), complete (C) and duplicated (D, Black), fragmented (F, Yellow), and missing (M, Red).

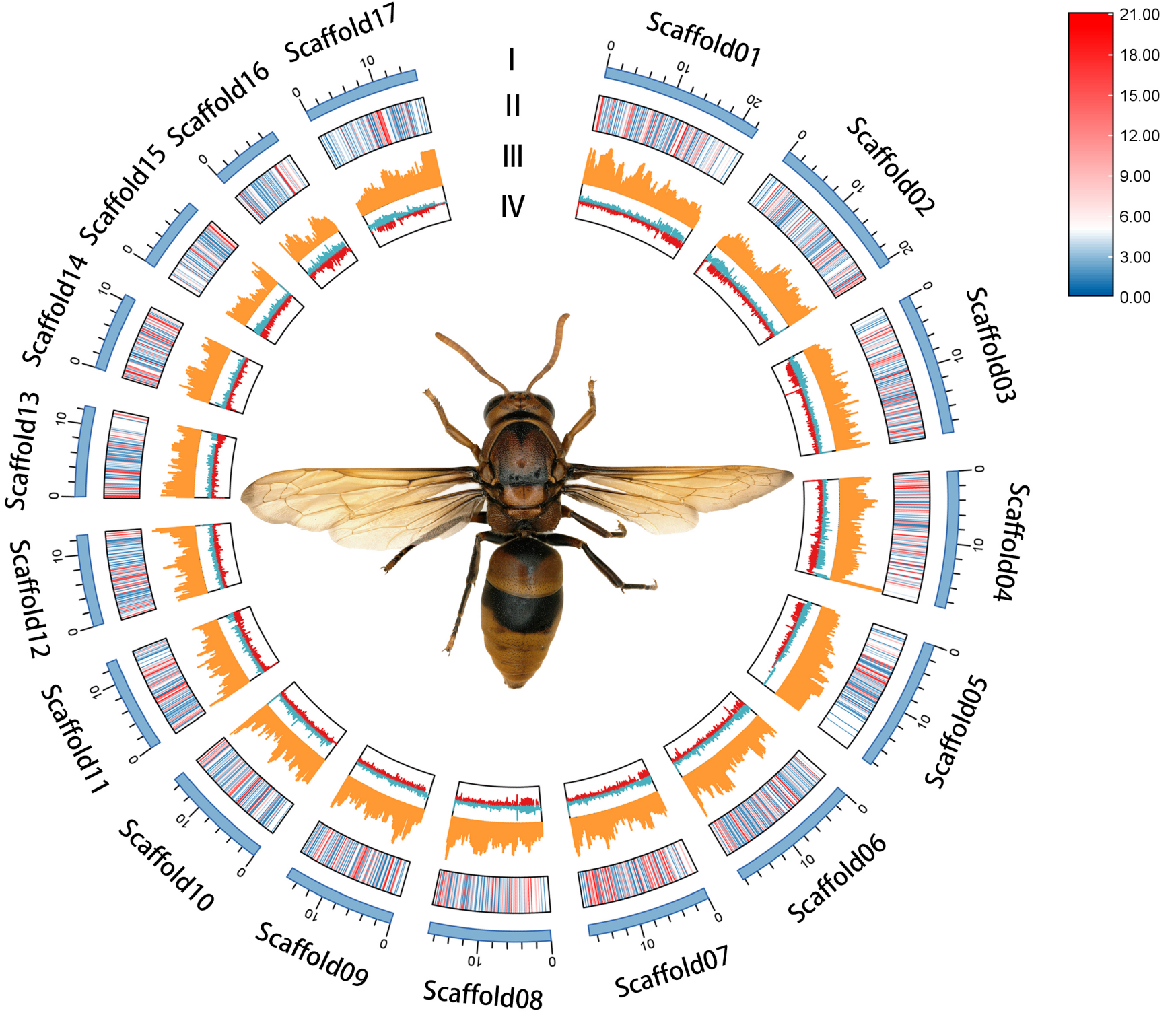
Fig. 4 Phylogenetic and evolutionary analyses of the *R. brunneum* genome. The results of each row in (**b**) and (**c**) correspond to the species name in the same row in (**a**). (**a**) Phylogenetic relationships and divergence times are shown for *R. brunneum* with other hymenopterans; (**b**) Expansions and contractions of gene families. Blue and yellow indicate the expanded and contracted gene families, respectively; (**c**) Distribution of genes in different hymenopteran species. The colored histogram indicates that genes of each species were categorized into five groups: Universal single-copy genes (single copy orthologous genes in common gene families); Multi-copy genes (multiple copy orthologous genes in common gene families); Species-specific (genes from unique gene families from each species); Others (genes that do not belong to any of the above ortholog categories); Unassigned (genes which are not clustered into any family).

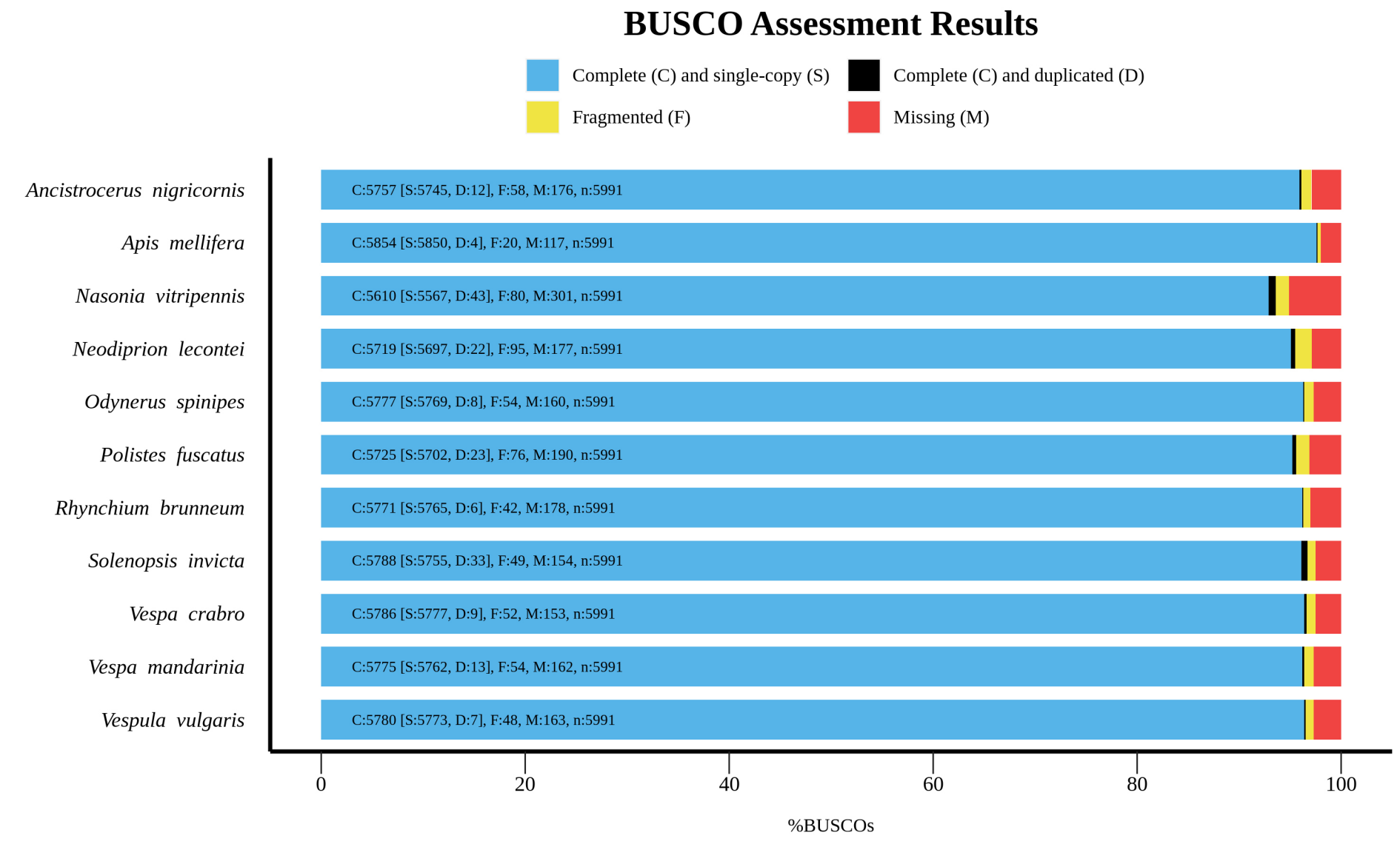
Fig. 5 (**a**) GO enrichment analyses of *R. brunneum* expanded gene families; (**b**) KEGG enrichment analyses of *R. brunneum* expanded gene families; (**c**) Chromosomal synteny between *V. crabro*, *R. brunneum* and *A. mellifera* genomes.

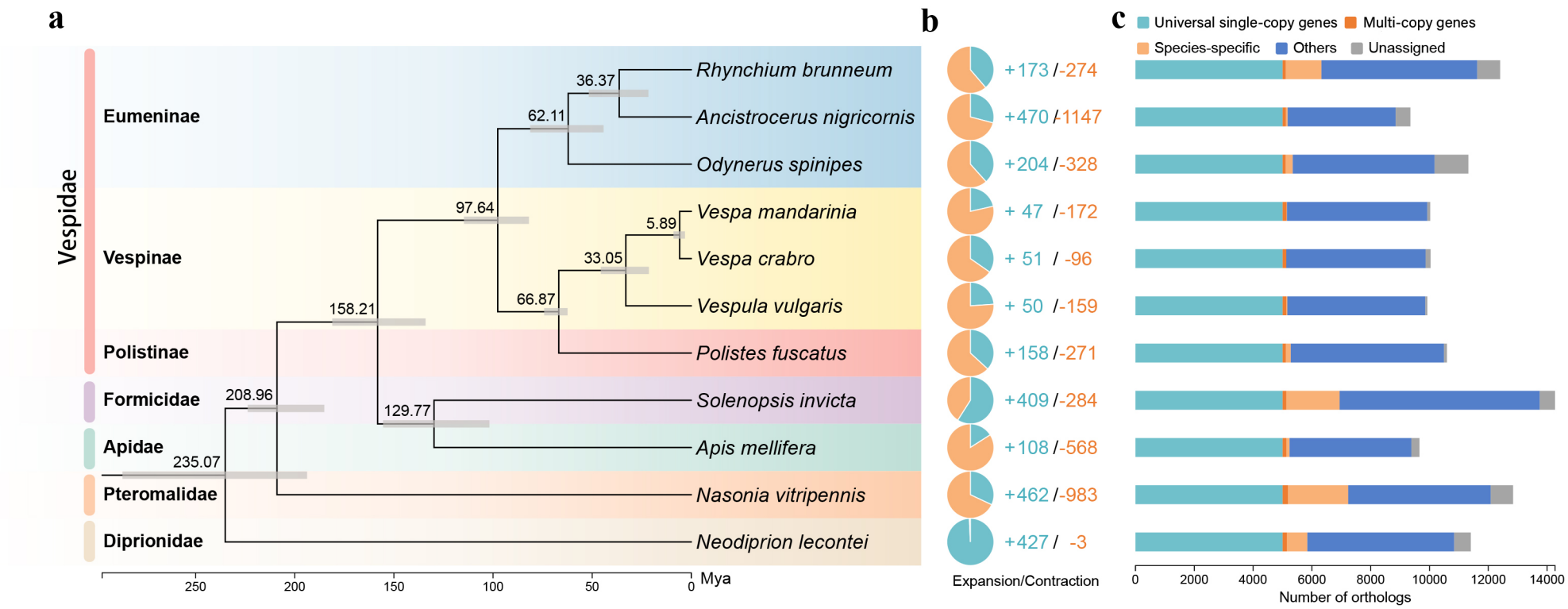

a
SSC-H
FSC-H
b
SSC-H
FL2-H
c
cell counts
fluorescence pulse height

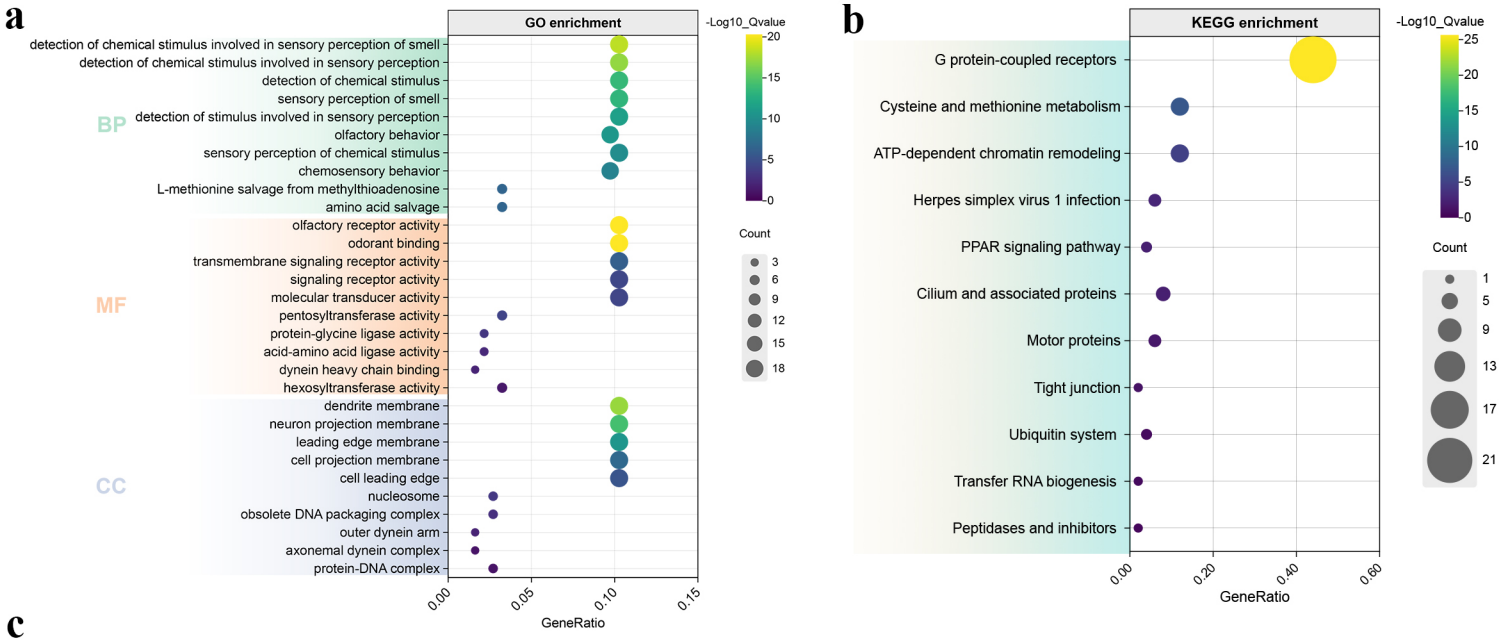

d
proposed diploid
log kmers pairs
493000
55000
6000
700
80
10
0
AB 0.97
AAB 0.03
AAB
AB
1n = 57
Total coverage of the kmer pair: A + B
Normalized minor kmer coverage: B / (A + B)
1/5
1/4
1/3
2/5
1/2

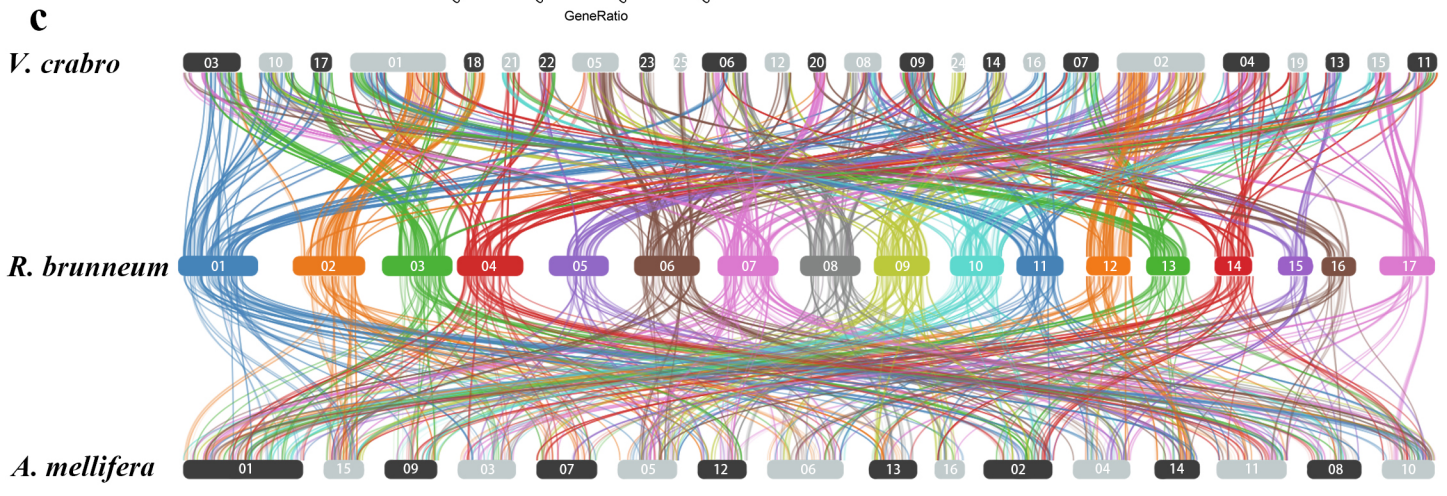

e
0 MB
100 MB
200 MB


Scaffold01
Scaffold02
Scaffold03
Scaffold04
Scaffold05
Scaffold06
Scaffold07
Scaffold08
Scaffold09
Scaffold10
Scaffold11
Scaffold12
Scaffold13
Scaffold14
Scaffold15
Scaffold16
Scaffold17
I
II
III
IV
21.00
18.00
15.00
12.00
9.00
6.00
3.00
0.00


BUSCO Assessment Results
Complete (C) and single-copy (S)
Complete (C) and duplicated (D)
Fragmented (F)
Missing (M)
Ancistrocerus nigricornis
C:5757 [S:5745, D:12], F:58, M:176, n:5991
Apis mellifera
C:5854 [S:5850, D:4], F:20, M:117, n:5991
Nasonia vitripennis
C:5610 [S:5567, D:43], F:80, M:301, n:5991
Neodiprion lecontei
C:5719 [S:5697, D:22], F:95, M:177, n:5991
Odynerus spinipes
C:5777 [S:5769, D:8], F:54, M:160, n:5991
Polistes fuscatus
C:5725 [S:5702, D:23], F:76, M:190, n:5991
Rhynchium brunneum
C:5771 [S:5765, D:6], F:42, M:178, n:5991
Solenopsis invicta
C:5788 [S:5755, D:33], F:49, M:154, n:5991
Vespa crabro
C:5786 [S:5777, D:9], F:52, M:153, n:5991
Vespa mandarinia
C:5775 [S:5762, D:13], F:54, M:162, n:5991
Vespula vulgaris
C:5780 [S:5773, D:7], F:48, M:163, n:5991
0
20
40
60
80
100
%BUSCOs


a
Vespidae
Eumeninae
Vespinae
Polistinae
Formicidae
Apidae
Pteromalidae
Diprionidae
Rhynchium brunneum
Ancistrocerus nigricornis
Odynerus spinipes
Vespa mandarinia
Vespa crabro
Vespula vulgaris
Polistes fuscatus
Solenopsis invicta
Apis mellifera
Nasonia vitripennis
Neodiprion lecontei
36.37
62.11
97.64
5.89
33.05
66.87
158.21
208.96
129.77
235.07
250
200
150
100
50
0
Mya
b
+173 /-274
+470 /-1147
+204 /-328
+ 47 /-172
+ 51 / -96
+ 50 /-159
+158 /-271
+409 /-284
+108 /-568
+462 /-983
+427 / -3
Expansion/Contraction
c
Universal single-copy genes
Multi-copy genes
Species-specific
Others
Unassigned
0
2000
4000
6000
8000
10000
12000
14000
Number of orthologs


a
GO enrichment
-Log10_Qvalue
Count
BP
MF
CC
detection of chemical stimulus involved in sensory perception of smell
detection of chemical stimulus involved in sensory perception
detection of chemical stimulus
sensory perception of smell
detection of stimulus involved in sensory perception
olfactory behavior
sensory perception of chemical stimulus
chemosensory behavior
L-methionine salvage from methylthioadenosine
amino acid salvage
olfactory receptor activity
odorant binding
transmembrane signaling receptor activity
signaling receptor activity
molecular transducer activity
pentosyltransferase activity
protein-glycine ligase activity
acid-amino acid ligase activity
dynein heavy chain binding
hexosyltransferase activity
dendrite membrane
neuron projection membrane
leading edge membrane
cell projection membrane
cell leading edge
nucleosome
obsolete DNA packaging complex
outer dynein arm
axonemal dynein complex
protein-DNA complex
GeneRatio
b
KEGG enrichment
-Log10_Qvalue
Count
G protein-coupled receptors
Cysteine and methionine metabolism
ATP-dependent chromatin remodeling
Herpes simplex virus 1 infection
PPAR signaling pathway
Cilium and associated proteins
Motor proteins
Tight junction
Ubiquitin system
Transfer RNA biogenesis
Peptidases and inhibitors
GeneRatio


c
V. crabro
R. brunneum
A. mellifera