scientific data

Check for updates

OPEN

DATA DESCRIPTOR

Chromosome-level genome assembly of *Trichogramma ostriniae* Pang & Chen (Hymenoptera: Trichogrammatidae)

Xue Sheng, Tian-Yi Yao, Yuan-Xi Li & Zhi-Chao Yan ✉

Trichogramma ostriniae **(Hymenoptera: Trichogrammatidae) is an important biological control agent that is effective against the Asian corn borer *Ostrinia furnacalis*. In this study, we present a chromosome-scale genome assembly of *T. ostriniae*. By integrating PacBio long-read sequencing and Hi-C scaffolding techniques, we constructed a genome assembly of 221.86 Mb. Approximately 93.14% of the assembly was successfully anchored onto 5 chromosomes, with contig and scaffold N50 values reaching 12.42 Mb and 39.71 Mb, respectively. The genome completeness, as evaluated by BUSCO, achieved a completeness score of 98.3%. Additionally, a total of 12,707 protein-coding genes were identified. In conclusion, this high-quality genome assembly of *T. ostriniae* serves as a valuable genomic resource, providing insights into the evolution of parasitoid diversity and enhancing the application of parasitoid wasps in biological control.**

Background & Summary

Parasitoid wasps, a group of parasitic insects with rich species diversity, have played pivotal roles in both biological control and evolutionary studies[1–7]. The Asian corn borer (ACB), *Ostrinia furnacalis*, is a destructive maize pest in Southeast Asia that causes substantial economic losses annually[8]. *Trichogramma dendrolimi* and *T. ostriniae* have been widely employed as effective biological control agents against ACB[9]. Extensive studies have demonstrated that, compared with *T. dendrolimi*, *T. ostriniae* has significantly superior pest control efficacy, highlighting its high value in practical biological control applications[9–13]. Furthermore, *T. ostriniae* serves as an ideal model for studying key traits in parasitoid wasps, such as reduced body size[3,14–16], trait loss[4,17,18], host range shift[19–22], and sex determination[23–26]. With the advancement of sequencing technologies, several *Trichogramma* species sequenced and their chromosome-level genome assemblies are publicly available, including *T. dendrolimi*[27], *T. chilonis*[28] and *T. japonicum*[29]. However, the lack of a high-quality genome for *T. ostriniae* has hindered in-depth research into its biological characteristics and applications in biological control.

Here, we present a high-quality, chromosome-level genome assembly of *T. ostriniae* using a multi-omics approach combining PacBio sequencing, Illumina sequencing, and chromatin conformation capture (Hi-C) techniques. Additionally, we integrated Oxford Nanopore Technologies (ONT) full-length transcriptome sequencing to further improve genome annotation quality. The high-quality genome of *T. ostriniae* provides a valuable resource for future studies of these diverse and economically important insects, including explorations of parasitoid biology, biological control, and the evolution of miniaturization.

Methods

Sample preparation and sequencing. The parasitoid wasps *T. ostriniae* used in this study were from Shanxi Province and were reared in our laboratory for more than a decade. The wasps were reared on the eggs of *Corcyra cephalonica* at 25 ± 1°C with 70 ± 5% relative humidity and a 14:10 h (light:dark) photoperiod. Adults were fed a 10% sucrose solution after emergence.

Female wasp genomic DNA was purified using a QIAGEN® Blood & Cell Culture DNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's procedure for both long-read and short-read whole-genome

State Key Laboratory of Agricultural and Forestry Biosecurity, College of Plant Protection, Nanjing Agricultural University, Nanjing, 210095, China. ✉e-mail: zcyan@njau.edu.cn

Sequencing strategy	Platform	Usage	Clean data (Gb)	Coverage (×)
Short-reads	Illumina	Genome Survey	8.36	36.71
Long-reads	PacBio	Assembly	11.43	51.52
Hi-C	Illumina	Hi-C assembly	56.75	274.66
RNA-seq	ONT	Annotation	85.01	—

Table 1. Statistics of the sequencing data used for genome assembly. ONT: Oxford Nanopore Technologies.

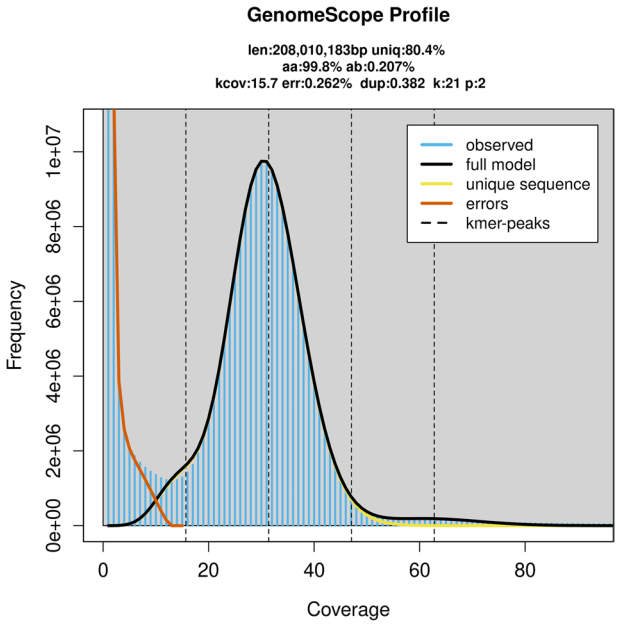

Fig. 1 GenomeScope profiles of 21-mer analysis.

sequencing. Total RNA was extracted with TRIzol (Invitrogen, California, CA, USA) according to the manufacturer's instructions. The concentration and purity of RNA were detected using NanoDrop One spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and RNA integrity was assessed using the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

Long-read genome sequencing was performed on the PacBio Sequel II platform. To anchor scaffolds to chromosomes, a high-throughput chromosome conformation capture (Hi-C) library was prepared using established methods[30] and sequenced on the Illumina NovaSeq X Plus platform. Illumina whole-genome sequencing (WGS) was performed on the Illumina NovaSeq 6000 platform using a 150 bp paired-end strategy. Full-length transcriptome sequencing was performed using the PromethION platform from Oxford Nanopore Technologies (ONT). Finally, our genome sequencing yielded 11.43 Gb (51.52×) of PacBio long reads, 8.36 Gb (36.71×) of Illumina short reads, 56.75 Gb (274.66×) of Hi-C data, and 85.01 Gb of transcriptome data (Table 1).

Genome survey and assembly. Prior to genome assembly, a k-mer analysis was conducted using Illumina sequencing data to infer the genomic characteristics, including genome size and heterozygosity. Briefly, quality-filtered reads were subjected to a 21-mer frequency distribution analysis employing Jellyfish v2.2.10[31]. Furthermore, the overall genomic properties were inferred via GenomeScope v2.0[32]. A preliminary genome survey of *T. ostriniae* revealed an estimated genome size of 208.01 Mb, with a heterozygosity of 0.207% (Fig. 1).

De novo genome assembly was performed using Hifiasm v0.13[33] with the PacBio HiFi reads. To eliminate potential contaminants, the resulting contigs were aligned against the NCBI nucleotide database via BLASTN with an E-value threshold of $1e^{-3}$, followed by the removal of identified contaminating contigs and mitochondrial contigs. The decontaminated contigs were subsequently scaffolded using Hi-C data. Specifically, Hi-C reads were aligned to the decontaminated contigs using Juicer v1.6[34] with the restriction enzyme MboI. Chromosome-scale scaffolding was performed using 3D-DNA v180114[35]. Any placement and orientation errors that showed clear discrete chromatin interaction patterns were manually corrected and visualized using Juicebox v1.11.08[36]. Finally, the quality of the final assembly was assessed using BUSCO v5.2.2[37] with the insecta_odb10 dataset. The final assembly size was 221.86 Mb, which comprised 206.63 Mb of sequences anchored into five chromosomes and 15.23 Mb of unplaced scaffolds. The assembly exhibited high continuity, with a scaffold N50 of 39.71 Mb and the longest scaffold being 55.07 Mb. The assembly was scaffolded into 5 chromosomes. The chromosomal sizes ranged from 33.67 Mb to 55.07 Mb. (Figs. 2, 3; Tables 2, 3).

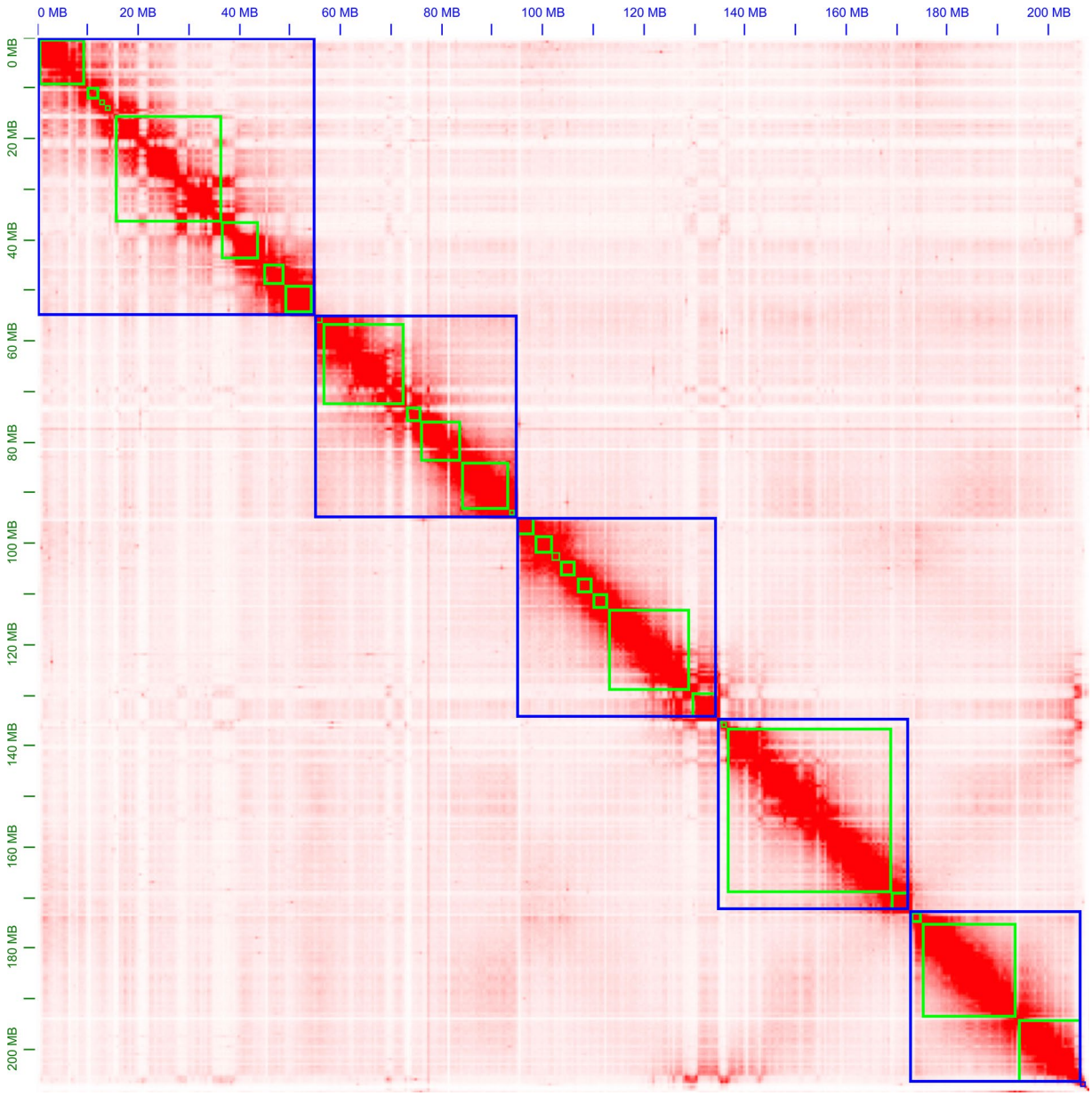

Fig. 2 Genome-wide chromosomal heatmap of *Trichogramma ostriniae*, where each chromosome is delineated in blue and each contig is highlighted with a green outline.

Repeat and non-coding RNA annotation. For the *T. ostriniae* genome, transposon element (TE) identification was performed using the Extensive *de novo* TE Annotator (EDTA) v1.9.6[38]. A custom repeat library was constructed by integrating known repeats from the RepBase-20180826[39] and Dfam[40] databases with the *de novo* TE detection results. RepeatMasker v4.1.2[41] was then used to mask repetitive elements in the *T. ostriniae* genome, revealing a total of 61.91 Mb of repetitive sequences, accounting for 27.90% of the genome (Table 4).

Additionally, transfer RNAs and ribosomal RNAs were detected using tRNAscan-SE v2.0.9[42] and RNAmmer v1.2[43], respectively. In total, 1513 ncRNAs were predicted, including 781 ribosomal RNAs (rRNAs) and 732 transfer RNAs (tRNAs) (Table 4).

Gene prediction and functional annotation. Gene prediction was performed through an integrative strategy combining *ab initio*, homology-based, and long-read transcriptome-guided approaches. *De novo* gene prediction of the softmasked genome was performed using Augustus v3.4.0[44]. For homology-based gene prediction, invertebrate protein sequences downloaded from the NCBI RefSeq database served as a reference dataset. GeneWise v2.4.1[45] was utilized to predict gene structures by aligning these reference proteins to the genomic sequence. For transcriptome-based analysis, Nanopore cDNA sequences were first oriented and trimmed using Pychopper v2.5.0 and then mapped to the assembled genome via Minimap2 v2.22[46]. StringTie v2.1.7[47] was subsequently employed to obtain transcript locations, followed by open reading frame prediction via TransDecoder

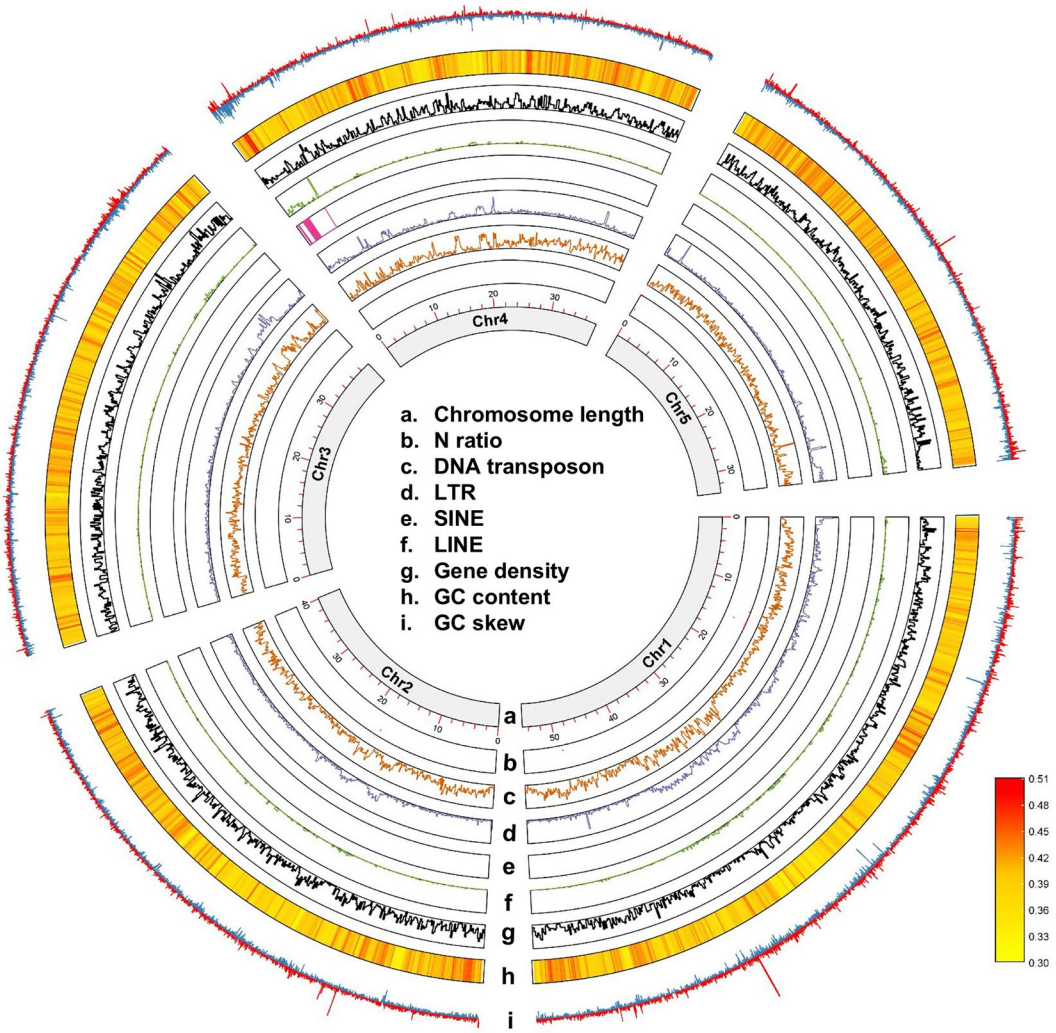

Fig. 3 Circos plot of the distribution of the genomic elements in *Trichogramma ostriniae*. The tracks indicate (**a**) the length of the chromosome (Mb), (**b**) the proportion of unknown bases (N), (**c**) DNA transposons, (**d**) long terminal repeats (LTR), (**e**) short interspersed nuclear elements (SINE), (**f**) long interspersed nuclear elements (LINE), (**g**) gene density, (**h**) GC content and (**i**) GC skew. The densities of TEs, genes, and GCs were calculated in 100 kb windows.

Assembly size (Mb)	Number of scaffolds/contigs (chromosomes)	Scaffold/contig N50 length (Mb)	GC (%)	BUSCO (n = 1,367) (%)			
				C	D	F	M
221.86	236/353 (5)	39.71/12.42	39.78	98.3	4.4	0.1	1.6

Table 2. Statistics of the final genome assembly of *T. ostriniae*. C: complete BUSCOs; D: complete and duplicated BUSCOs; F: fragmented BUSCOs; M: missing BUSCOs.

Chromosome	Length (Mb)
Chr01	55.07
Chr02	40.19
Chr03	39.71
Chr04	37.99
Chr05	33.67
Total	206.63
Unplaced	15.23 (6.86%)

Table 3. Statistics for chromosome sequence length.

Repetitive elements	
Size (Mb)	61.91 (27.90%)
DNA transposons (Mb)	29.61 (13.35%)
LTRs (Mb)	18.45 (8.32%)
LINEs (Kb)	970.93 (0.44%)
SINEs (bp)	1296 (0.00%)
Unclassified (Mb)	12.87 (5.80%)
ncRNAs	
tRNA	732
rRNA	781
Protein-coding genes	
Total number	12,707
BUSCO completeness (%)	94.4

Table 4. Genome annotation statistics of *T. ostriniae*.

v5.5.0[48]. The resulting protein sequences were aligned against the Swiss-Prot database using BLASTP, with an E value threshold of 1e-5. EVidenceModeler (EVM) v1.1.1[49] was used to integrate the genes predicted by the three methods to form a consensus gene set. The EVM results were iteratively refined through two rounds of updates using Program to Assemble Spliced Alignments (PASA) v2.4.1[49] to generate an optimized gene set (OGS). By combining *de novo*, homology-based and transcriptome-based predictions, a total of 14,894 protein-coding genes were predicted in the *T. ostriniae* genome. The completeness of the protein sequences was assessed using BUSCO on the basis of the OrthoDB database insecta_odb10[50]. The completeness of BUSCO genes is approximately 94.4%, with 90.2% complete single-copy, 4.2% complete duplicated, 1.2% fragmented, and 4.4% missing (Table 4).

Functional annotation was performed by aligning the predicted protein sequences to the non-redundant protein (Nr) database and SwissProt database, using Diamond v2.0.11[51] with an E-value threshold of $1e^{-5}$. InterProScan 5.53-87.0[52] and eggNOGmapper v2.0.1[53] were used to annotate the Gene Ontology (GO) terms and pathways (Kyoto Encyclopedia of Genes and Genomes, Reactome). In total, 12707 genes were functionally annotated, representing 85.32% of the total predicted protein-coding genes (Table 4).

Data Records

The whole-genome project is accessible in the NCBI database under BioProject accession number PRJNA1313832. The Illumina, PacBio, Hi-C and RNA-seq data are available under accession numbers SRP616411[54]. The assembled genome of *T. ostriniae* is available in GenBank under accession number JBQXEY000000000.1[55]. Genome annotations are available in the FigShare repository[56].

Technical Validation

The integrity of the DNA samples was assessed via 0.75% agarose gel electrophoresis to confirm their integrity. The purity of the extracted DNA was subsequently evaluated via a NanoDrop One spectrophotometer (Thermo Fisher Scientific, USA), with acceptable OD260/280 ratios ranging from 1.8 to 2.0 and OD260/230 ratios ranging from 2.0 to 2.2. The precise concentration of DNA was determined via a Qubit 3.0 fluorometer (Life Technologies, CA, USA). The degradation and contamination of RNA were examined via 1% agarose gel electrophoresis. The RNA concentration was quantified via a Qubit RNA Assay Kit in conjunction with a Qubit Fluorometer. Additionally, RNA integrity was assessed via an RNA Nano 6000 Assay Kit on a Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

The assembled genome size is 221.86 Mb, with repetitive sequences accounting for 27.90% of the total. The size of the assembled genome is consistent with the estimated result of 208.01 Mb and is similar to the genome sizes of *T. chilonis*[28], *T. dendrolimi*[27], and *T. japonicum*[29]. We evaluated assembly completeness utilizing BUSCO v5.2.2[37] against the Insecta reference gene set (n = 1,367). The assembly exhibited a BUSCO completeness of 98.3% BUSCO genes, comprising 4.4% duplicated, 0.1% fragmented, and 1.6% missing genes. These results demonstrate that we successfully obtained a high-quality genome for *T. ostriniae*.

Our integrated annotation approach successfully identified 12,707 protein-coding genes, achieving exceptional genome annotation quality, as evidenced by the high BUSCO completeness score (94.4%). Combining *de novo*, transcriptomic, and homology-based methods enabled comprehensive gene detection, with functional annotations assigned to 85.32% of the predicted genes. The 14.7% of genes lacking functional annotation likely include important lineage-specific adaptation elements, highlighting valuable targets for future functional characterization. This high-quality annotation supports advanced genomic research in parasitoid wasps.

Data availability

The genome assembly and annotation datasets are available in the FigShare repository under accession code https://doi.org/10.6084/m9.figshare.30073240.v2. The raw sequencing reads have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession PRJNA1313832. The assembled genome is available under GenBank accession JBQXEY000000000.

Code availability

No custom scripts or codes were used in this study. All bioinformatics analyses were conducted using standard protocols as outlined in the documentation of the respective tools and software. The software versions and parameters used are detailed throughout the manuscript.

Received: 22 September 2025; Accepted: 9 December 2025;
Published online: 19 December 2025

References

1. Burke, G. R. & Sharanowski, B. J. Parasitoid wasps. *Curr. Biol.* **34**, R483–R488, https://doi.org/10.1016/j.cub.2024.03.038 (2024).
2. Peters, R. S. *et al*. Evolutionary History of the Hymenoptera. *Curr. Biol.* **27**, 1013–1018, https://doi.org/10.1016/j.cub.2017.01.027 (2017).
3. Lindsey, A. R. I. *et al*. Comparative genomics of the miniature wasp and pest control agent *Trichogramma pretiosum*. *BMC Biol.* **16**, 54, https://doi.org/10.1186/s12915-018-0520-9 (2018).
4. Zhao, X. *et al*. Genomic signatures associated with the evolutionary loss of egg yolk in parasitoid wasps. *Proc. Natl. Acad. Sci. USA.* **122**, e2422292122, https://doi.org/10.1073/pnas.2422292122 (2025).
5. Huang, J. *et al*. Two novel venom proteins underlie divergent parasitic strategies between a generalist and a specialist parasite. *Nat Commun* **12**, 234, https://doi.org/10.1038/s41467-020-20332-8 (2021).
6. Smith, S. M. Biological control with *Trichogramma*: advances, successes, and potential of their use. *Annu. Rev. Entomol.* **41**, 375–406, https://doi.org/10.1146/annurev.en.41.010196.002111 (1996).
7. Tougeron, K. *et al*. Ecology and biology of the parasitoid *Trechnites insidiosus* and its potential for biological control of pear psyllids. *Pest Manag Sci* **77**, 4836–4847, https://doi.org/10.1002/ps.6517 (2021).
8. Abbas, A. *et al*. Global distribution and sustainable management of Asian corn borer (ACB), *Ostrinia furnacalis* (Lepidoptera: Crambidae): recent advancement and future prospects. *Bull. Entomol. Res.* **115**, 105–120, https://doi.org/10.1017/S0007485324000919 (2025).
9. Zang, L. S., Wang, S., Zhang, F. & Desneux, N. Biological Control with *Trichogramma* in China: History, Present Status, and Perspectives. *Annu. Rev. Entomol.* **66**, 463–484, https://doi.org/10.1146/annurev-ento-060120-091620 (2021).
10. Hoffmann, M. P., Walker, D. L. & Shelton, A. M. Biology of *Trichogramma ostriniae* (Hym.: trichogrammatidae) reared on *Ostrinia nubilalis* (Lep.: pyralidae) and survey for additional hosts. *Entomophaga* **40**, 387–402, https://doi.org/10.1007/BF02373727 (1995).
11. Wang, Y. *et al*. Insecticide toxic effects on *Trichogramma ostriniae* (Hymenoptera: Trichogrammatidae). *Pest Manag Sci* **68**, 1564–1571, https://doi.org/10.1002/ps.3343 (2012).
12. Wang, Z.-Y., He, K.-L., Zhang, F., Lu, X. & Babendreier, D. Mass rearing and release of *Trichogramma* for biological control of insect pests of corn in China. *Biol. Control* **68**, 136–144, https://doi.org/10.1016/j.biocontrol.2013.06.015 (2014).
13. Wang, Y. *et al*. *Trichogramma ostriniae* is more effective than *Trichogramma dendrolimi* as a biocontrol agent of the Asian Corn Borer, *Ostrinia furnacalis*. *Insects* **13**, 70, https://doi.org/10.3390/insects13010070 (2022).
14. Xu, H. *et al*. Comparative Genomics Sheds Light on the Convergent Evolution of Miniaturized Wasps. *Mol. Biol. Evol.* **38**, 5539–5554, https://doi.org/10.1093/molbev/msab273 (2021).
15. Greenberg, S. M., Nordlund, D. A. & Wu, Z. Influence of Rearing Host on Adult Size and Ovipositional Behavior of Mass Produced Female *Trichogramma minutumRiley* and *Trichogramma pretiosumRiley* (Hymenoptera: Trichogrammatidae). *Biol. Control* **11**, 43–48, https://doi.org/10.1006/bcon.1997.0582 (1998).
16. Bai, B., Luck, R. F., Forster, L., Stephens, B. & Janssen, J. A. M. The effect of host size on quality attributes of the egg parasitoid, *Trichogramma pretiosum*. *Entomol. Exp. Appl.* **64**, 37–48, https://doi.org/10.1111/j.1570-7458.1992.tb01592.x (1992).
17. Lammers, M., Kraaijeveld, K., Mariën, J. & Ellers, J. Gene expression changes associated with the evolutionary loss of a metabolic trait: lack of lipogenesis in parasitoids. *BMC Genomics* **20**, 309, https://doi.org/10.1186/s12864-019-5673-6 (2019).
18. Visser, B. *et al*. Phenotypic plasticity explains apparent reverse evolution of fat synthesis in parasitic wasps. *Sci. Rep.* **11**, 7751, https://doi.org/10.1038/s41598-021-86736-8 (2021).
19. Pennacchio, F. & Strand, M. R. Evolution of developmental strategies in parasitic Hymenoptera. *Annu. Rev. Entomol.* **51**, 233–258, https://doi.org/10.1146/annurev.ento.51.110104.151029 (2006).
20. Ma, Y. *et al*. The genome of the rice planthopper egg parasitoid wasps *Anagrus nilaparvatae* casts light on the chemo- and mechanosensation in parasitism. *BMC Genomics* **23**, 541, https://doi.org/10.1186/s12864-022-08656-9 (2022).
21. Robertson, H. M., Gadau, J. & Wanner, K. W. The insect chemoreceptor superfamily of the parasitoid jewel wasp *Nasonia vitripennis*. *Insect Mol. Biol.* **19**, 121–136, https://doi.org/10.1111/j.1365-2583.2009.00979.x (2010).
22. Heimpel, G. E., Abram, P. K. & Brodeur, J. A phylogenetic perspective on parasitoid host ranges with implications for biological control. *Curr Opin in Insect Sci* **44**, 95–100, https://doi.org/10.1016/j.cois.2021.04.003 (2021).
23. Heimpel, G. E. & de Boer, J. G. Sex Determination in the Hymenoptera. *Annu. Rev. Entomol.* **53**, 209–230, https://doi.org/10.1146/annurev.ento.53.103106.093441 (2008).
24. Li, C. *et al*. *Wolbachia* symbionts control sex in a parasitoid wasp using a horizontally acquired gene. *Curr. Biol.* **34**, 2359–2372. e2359, https://doi.org/10.1016/j.cub.2024.04.035 (2024).
25. Zou, Y., Geuverink, E., Beukeboom, L. W., Verhulst, E. C. & van de Zande, L. A chimeric gene paternally instructs female sex determination in the haplodiploid wasp *Nasonia*. *Science* **370**, 1115–1118, https://doi.org/10.1126/science.abb8949 (2020).
26. Stouthamer, R., Breeuwer, J. A. J., Luck, R. F. & Werren, J. H. Molecular identification of microorganisms associated with parthenogenesis. *Nature* **361**, 66–68, https://doi.org/10.1038/361066a0 (1993).
27. Dong, Q. *et al*. Chromosomal-level genome assembly of *Trichogramma dendrolimi* (Trichogrammatidae: Hymenoptera). *Sci. Data* **12**, 667, https://doi.org/10.1038/s41597-025-04997-8 (2025).
28. Wang, C. *et al*. A chromosomal-level genome assembly of *Trichogramma chilonis* Ishii, 1941 (Hymenoptera: Trichogrammatidae). *Sci. Data* **12**, 457, https://doi.org/10.1038/s41597-025-04792-5 (2025).
29. Jiang, Z. *et al*. Chromosomal-level genome assembly of *Trichogramma japonicum* (Hymenoptera: Trichogrammatidae). *Sci. Data* **12**, 1537, https://doi.org/10.1038/s41597-025-05910-z (2025).
30. Belton, J. M. *et al*. Hi-C: A comprehensive technique to capture the conformation of genomes. *Methods* **58**, 268–276, https://doi.org/10.1016/j.ymeth.2012.05.001 (2012).
31. Marçais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27**, 764–770, https://doi.org/10.1093/bioinformatics/btr011 (2011).
32. Ranallo-Benavidez, T. R., Jaron, K. S. & Schatz, M. C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat Commun* **11**, 1432, https://doi.org/10.1038/s41467-020-14998-3 (2020).
33. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175, https://doi.org/10.1038/s41592-020-01056-5 (2021).
34. Durand, N. C. *et al*. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst* **3**, 95–98, https://doi.org/10.1016/j.cels.2016.07.002 (2016).
35. Zhang, Y., Xiong, Y. & Xiao, Y. 3dDNA: A Computational Method of Building DNA 3D Structures. *Molecules* **27**, 5936, https://doi.org/10.3390/molecules27185936 (2022).

36. Robinson, J. T. *et al.* Juicebox.js Provides a Cloud-Based Visualization System for Hi-C Data. *Cell Syst* **6**, 256–258.e251, https://doi.org/10.1016/j.cels.2018.01.001 (2018).
37. Seppey, M., Manni, M. & Zdobnov, E. M. BUSCO: Assessing Genome Assembly and Annotation Completeness. *Methods Mol. Biol.*, 227–245, https://doi.org/10.1007/978-1-4939-9173-0_14 (2019).
38. Su, W., Ou, S., Hufford, M. B. & Peterson, T. A Tutorial of EDTA: Extensive *De Novo* TE Annotator. *Methods Mol. Biol.*, 55–67, https://doi.org/10.1007/978-1-0716-1134-0_4 (2021).
39. Bao, W., Kojima, K. K. & Kohany, O. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA* **6**, 11, https://doi.org/10.1186/s13100-015-0041-9 (2015).
40. Hubley, R. *et al.* The Dfam database of repetitive DNA families. *Nucleic Acids Res.* **44**, D81–D89, https://doi.org/10.1093/nar/gkv1272 (2015).
41. Tarailo-Graovac, M. & Chen, N. Using RepeatMasker to Identify Repetitive Elements in Genomic Sequences. *Curr Protoc Bioinformatics* **25**, 4.10.11–14.10.14, https://doi.org/10.1002/0471250953.bi0410s25 (2009).
42. Chan, P. P., Lin, B. Y., Mak, A. J. & Lowe, T. M. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res.* **49**, 9077–9096, https://doi.org/10.1093/nar/gkab688 (2021).
43. Lagesen, K. *et al.* RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **35**, 3100–3108, https://doi.org/10.1093/nar/gkm160 (2007).
44. Brůna, T., Hoff, K. J., Lomsadze, A., Stanke, M. & Borodovsky, M. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genom and Bioinform* **3**, https://doi.org/10.1093/nargab/lqaa108 (2021).
45. Birney, E., Clamp, M. & Durbin, R. GeneWise and Genomewise. *Comparative Study* **14**, 988–995, https://doi.org/10.1101/gr.1865504 (2004).
46. Li, H. Minimap and miniasm: fast mapping and *de novo* assembly for noisy long sequences. *Bioinformatics* **32**, 2103–2110, https://doi.org/10.1093/bioinformatics/btw152 (2016).
47. Kovaka, S. *et al.* Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* **20**, 278, https://doi.org/10.1186/s13059-019-1910-1 (2019).
48. Haas, B. J. *et al.* *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512, https://doi.org/10.1038/nprot.2013.084 (2013).
49. Haas, B. J. *et al.* Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* **9**, R7, https://doi.org/10.1186/gb-2008-9-1-r7 (2008).
50. Kriventseva, E. V. *et al.* OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Res.* **47**, D807–D811, https://doi.org/10.1093/nar/gky1053 (2018).
51. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60, https://doi.org/10.1038/nmeth.3176 (2015).
52. Finn, R. D. *et al.* InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res.* **45**, D190–D199, https://doi.org/10.1093/nar/gkw1107 (2016).
53. Huerta-Cepas, J. *et al.* Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Mol. Biol. Evol.* **34**, 2115–2122, https://doi.org/10.1093/molbev/msx148 (2017).
54. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRP616411 (2025).
55. *NCBI GenBank* https://identifiers.org/ncbi/insdc:JBQXEY000000000 (2025).
56. Sheng, X. Genome annotation of *Trichogramma ostriniae*. *figshare* https://doi.org/10.6084/m9.figshare.30073240.v2 (2025).

Acknowledgements

This study was supported by the National Key R & D Program of China (2024YFD1401200), the National Natural Science Foundation of China (32172483), Fundamental Research Funds for the Central Universities (Grant no. ZJ25195020), and the Natural Science Foundation of Hainan Province (Grant no. 323QN262). Bioinformatic analyses were supported by the high-performance computing platform of the Bioinformatics Center, Nanjing Agricultural University. Experiments were supported by Modern Agricultural Analysis and Testing Center, Nanjing Agricultural University.

Author contributions

Z.Y. and Y.L. contributed to the research design. T.Y. collected the samples. Z.Y. and X.S. analyzed the data. X.S., Y.L. and Z.Y. wrote the draft manuscript and revised it. All co-authors contributed to this manuscript and approved it.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Z.-C.Y.