scientific data

Check for updates

OPEN

DATA DESCRIPTOR

A chromosomal-level genome assembly of *Thanatophilus sinuatus* Fabricius (Coleoptera: Staphylinidae)

Yingna Zhang[1,2], Liangliang Li[3], Shipeng Shao[4], Jianfeng Jin[5], Yu Wang[6] ✉ & Ruixin Zhu[1] ✉

Thanatophilus sinuatus **(Coleoptera: Staphylinidae), a widespread necrophagous beetle frequently recovered from remains, is critical for post-mortem interval estimation. The absence of a reference genome has limited molecular studies of its forensic traits. Here, we report the first high-quality, chromosome-level genome assembly of *T. sinuatus*, built using PacBio HiFi, Illumina, and Hi-C data. The 288.95 Mb genome has a scaffold N50 of 18.97 Mb, with 98.37% of sequence (284.24 Mb) anchored to 16 chromosomes. BUSCO analysis (Insecta, n = 1,367) reveals 98.6% completeness with 97.3% single-copy BUSCOs and 1.3% duplicated BUSCOs. Repetitive elements comprise 34.32% of the assembly, and 12,300 protein-coding genes were predicted. This high-quality genome provides a valuable genomic resource for elucidating Silphinae comparative genomics, the genetic mechanisms underlying traits of forensic interest.**

Background & Summary

Thanatophilus sinuatus (Fabricius 1775) (Coleoptera: Staphylinidae) is a diurnal carrion beetle in Silphinae[1]. Adults are black, dorsoventrally flattened, with a sculptured pronotum, ribbed elytra, and sexually dimorphic elytral apices. *Thanatophilus sinuatus*, widespread in Holarctic and Afrotropical realms[2,3], is among the most abundant temperate carrion beetles[4], colonizing meadows, grasslands, montane slopes, forests, and agricultural habitats, underscoring its generalist habitat preference[1,5].

Thanatophilus sinuatus, the predominant necrophagous Silphinae beetle[6], could colonize corpses from early decomposition onward, thereby enhancing post-mortem interval (PMI) estimation and forensic utility[7]. Necrophagous insects perceive volatile organic compounds (VOCs) through olfaction and exhibit appropriate behavioral responses, such as carcasses localization, feeding preferences, mate selection, oviposition site selection, and breeding-site localization[8–15]. Studies of *T. sinuatus* immature morphology underpin precise instar and age determination[7,16,17], while controlled experiments on diet, constant and fluctuating temperatures, photoperiod, population density, and geography elucidate their impacts on development, survival, and behavior[2,4,17,18], collectively refining PMI estimates.

Carrion beetles were traditionally classified as a separate family Silphidae[7,19]. Molecular phylogenetics now place Silphidae as a derived clade within Staphylinidae[20–25]. So *T. sinuatus* falls in Staphylinidae: Silphinae: Silphini: Silphina[20,25,26]. This revision indicates that large necrophagous silphines evolved from small rove-beetle ancestors, reflecting a major niche expansion[20,25]. Phylogenetic trees retain two monophyletic carrion-beetle lineages—Silphini and Nicrophorini—as sister groups closest to Tachyporinae, supporting a single origin of necrophagy in their common ancestor[20,27]. *Thanatophilus* clusters with other wide-bodied silphines (*Silpha*,

[1]Shanghai Key Laboratory of Maternal Fetal Medicine, Shanghai Institute of Maternal-Fetal Medicine and Gynecologic Oncology, Clinical and Translation Research Center, Shanghai First Maternity and Infant Hospital, School of Life Sciences and Technology, Tongji University, Shanghai, 200092, P. R. China. [2]Department of Anatomy, Shihezi University School of Medicine, Shihezi, Xinjiang, China. [3]Shandong University of Political Science and Law, Jiefang East Road, Jinan, China. [4]The Judicial Appraisal Center of Soochow University, No. 178 Ganjiang East Road, Gusu District, Suzhou, China. [5]College of Life Sciences, Xinyang Normal University, Xinyang, 464000, China. [6]Department of Forensic Medicine, Soochow University, Ganjiang East Road, Suzhou, China. ✉e-mail: yuw@suda.edu.cn; rxzhu@tongji.edu.cn

Necrodes, *Oiceoptoma*) and appears to be a coherent lineage distinct from Nicrophorine, earlier taxonomy chiefly addressed regional synonyms[3,27].

Thanatophilus sinuatus and allied Silphinae illuminate the early evolution of carrion feeding in beetles. Most species feed on carrion and opportunistically on other insects[7,28]. Molecular evidence suggests that modern Silphinae evolved from rove-beetle ancestors that shifted to necrophagy early, with *T. sinuatus* exemplifying this shift[25]. *Thanatophilus sinuatus* can reach corpses rapidly and rivals fly larvae in decomposition[7], representing a separate evolutionary experiment in utilizing large vertebrate carcasses. Investigating *T. sinuatus* illuminates convergent necrophilic evolution and key carrion adaptations: refined chemosensation; predatory removal of fly larvae that mitigates competition; detoxification of carcass xenobiotics; and niche partitioning with *T. rugosus*, separated by colonisation sequence and carcass microhabitats[5,7,29]. *Thanatophilus sinuatus* is non-burying carrion beetle, unlike parental-care Nicrophorini. Silphini and Nicrophorini evolved contrasting carrion strategies: Silphini compete on exposed large carcasses, whereas Nicrophorini bury smaller ones. Molecular data place Silphinae's split from other Staphylinidae in the mid-Cretaceous and show genus-level diversification by the late Cretaceous–early Cenozoic[25]. *Thanatophilus sinuatus*, occupying a robust node within Silphini, epitomises this ancient, widely dispersed lineage, whose persistence underscores necrophagic resilience. Consequently, *Thanatophilus sinuatus* is a key model for carrion-driven evolution, including parental-care origins and detoxification adaptations.

Despite extensive ecological and forensic research on *T. sinuatus*, no chromosome-level genome assembly exists. Current data comprise mitochondrial COI, limited nuclear loci, and a few congeneric mitogenomes[25,30,31]. While informative for phylogenetics and identification, these markers cover only a tiny genomic fraction and lack information on nuclear genes and genomic architecture. Related taxa, such as Nicrophorus spp., have become emerging models for evolutionary genomics, with several species now having draft or chromosomal-level genomes. Notably, *Nicrophorus vespilloides* provided the first Silphinae genome, revealing DNA methylation and parental care-related genes[32,33], while the 202.3 Mb genome of *Nicrophorus investigator* was assembled, scaffolded into 7 chromosomal pseudomolecules, and annotated with 11,046 protein-coding genes, advancing Silphinae genomic studies[34].

The lack of a reference genome for *T. sinuatus*, a key Silphinae model, has hindered genetic and evolutionary research, making it difficult to address questions about gene content, function, and adaptive evolution. A high-quality genome assembly of *T. sinuatus* would reveal genomic signatures of its divergent lifestyle and facilitate the study of gene family evolution, focusing on expansions or contractions linked to its necrophagous behavior, with comparisons to other Coleoptera to identify lineage-specific innovations[35]. It would also aid in understanding the molecular basis of traits such as carrion chemosensory receptors (odorant receptors, gustatory receptors), digestive enzymes, immune responses in microbe-rich environments, and detoxification of carcass xenobiotics, including pathways like tyrosine metabolism[36], drug metabolism–cytochrome P450[37], and ascorbate–aldarate metabolism[38]. Additionally, a chromosome-level *T. sinuatus* genome enables fine-scale Silphinae phylogenetics[39], leveraging orthologous genes to reconstruct evolutionary history, sharpen Silphini–Nicrophorini placement, estimate divergence times, and illuminate carrion-beetle radiation and adaptation since the Mesozoic amid climatic and faunal shifts. Furthermore, a high-quality *T. sinuatus* genome will enable comparative genomics between burying and non-burying carrion beetles, refining Silphinae phylogeny and revealing genomic divergence among carrion-feeding strategies; chromosome-level comparisons with burying beetles will clarify structural shifts accompanying life-history divergence.

A chromosome-level *T. sinuatus* genome will accelerate forensic entomology by enabling microsatellite and SNP discovery for population genetics, rapid life-stage identification via barcoding or SNP assays, instar determination from expression profiles, and dissection of genes governing development, colonization timing, and thermal tolerance, thereby refining PMI estimates. Moreover, it will serve as a reference for metagenomic or environmental DNA surveys to detect insect evidence on remains. Genome-wide markers from our assembly will enable population genomics across the broad range of *T. sinuatus*, quantifying gene flow, regional genetic structure, and demographic history. Comparative analyses (e.g., Western Europe versus East Asia) can uncover migration routes, local adaptation to climate or habitat, and potential cryptic lineages, while also supporting the forensic geographic sourcing of insect evidence[3,5,17,40].

A chromosome-level assembly preserves synteny and long-range linkage, enabling robust analyses of karyotype evolution, chromosomal rearrangements, and gene-family expansions, while improving gene prediction, regulatory-element discovery, and trait-locus anchoring. For necrophagous beetles, such continuity facilitates detection of carrion-adapted loci under selection across species and strengthens phylogenetic inference by leveraging thousands of orthologues instead of sparse markers.

We propose a chromosome-level *T. sinuatus* genome assembly using PacBio HiFi, Illumina, and Hi-C data, annotated for repeats, non-coding RNAs, and protein-coding genes, advancing Silphinae comparative genomics, phylogenomics, ecology, evolution, and forensics.

Methods

Sample collection and preprocessing. Approximately 20 pairs of *T. sinuatus* were collected in October 2023 from Tangshan, Hebei, China (39.83°N, 118.16°E). A closed laboratory colony was established and maintained for at least 12 consecutive generations without immigration, resulting in a reduction of population heterozygosity due to the founder bottleneck followed by genetic drift. A single adult male *T. sinuatus* was used for both DNA (PacBio, Hi-C and Illumina) and RNA sequencing. Muscle tissue, specifically from the pronotum and posterior abdomen, was extracted from the specimen. The tissue was rinsed in phosphate-buffered saline for 10 min to remove external contaminants, flash-frozen in liquid nitrogen for 20 min, and stored at −80 °C until extraction and sequencing.

Libraries	Insert sizes (bp)	Clean data (Gb)	Sequencing coverage (×)
Illumina	350	29.83	103.23
PacBio Hi-Fi	15 Kb	19.74	60.31
Hi-C	350	45.70	158.16
RNA	350	9.68	—

Table 1. Statistics for the sequencing data used in the genome assembly of *Thanatophilus sinuatus*.

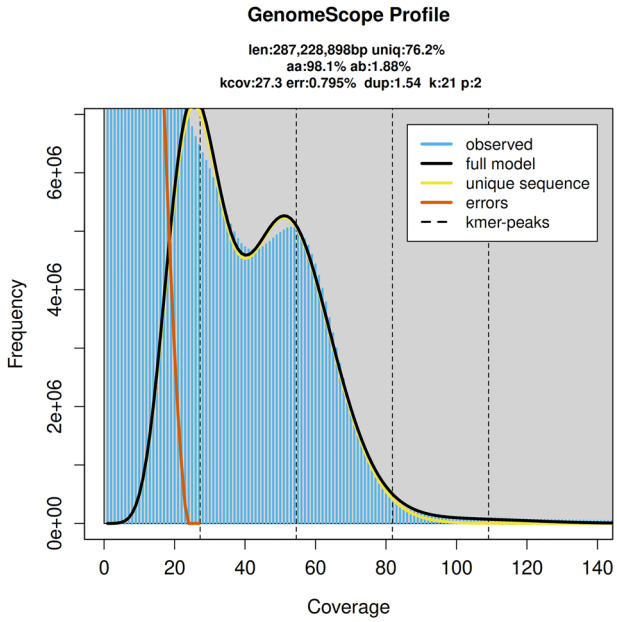

Fig. 1 Genomic survey of *Thanatophilus sinuatus* based on the 21-mer survey, performed using GenomeScope.

DNA extraction and library construction. Genomic DNA was extracted using the Cetyltrimethylammonium Bromide (CTAB) protocol and used for both Illumina next-generation sequencing and PacBio HiFi library construction.

The PacBio HiFi library (15 K insert) was constructed using the SMRTbell® Express Template Prep Kit 2.0 (Pacific Biosciences, PN 101-853-100) and run on the PacBio Sequel IIe platform. After quality control, DNA was sheared to 15 Kb using the Megaruptor system (Diagenode B06010001) and concentrated with AMPure® PB Beads (Pacific Biosciences 100-265-900). SMRTbell libraries were prepared following the protocol of the 2.0 kit, including steps for single-stranded overhang removal, DNA damage repair, end repair, A-tailing, adapter ligation, and enzymatic digestion. Fragment selection was performed using the SageELF (Sage Science ELF000) or BluePippin (Sage Science BLU0001) systems. For Illumina sequencing, short-read libraries (150 bp paired-end, 350 bp insert) were prepared with the TruSeq DNA PCR-free kit and sequenced on the Illumina NovaSeq 6000.

RNA extraction and library construction. RNA was extracted using the TRIzol™ Reagent kit (Thermo Fisher Scientific), and RNA libraries were constructed using the TruSeq RNA v2 kit, followed by sequencing on the Illumina NovaSeq 6000.

Hi-C library construction. Hi-C libraries were generated following the standard protocol[41]. Briefly, samples were cross-linked with 1% formaldehyde for 10 min at room temperature and quenched with 0.125 M glycine for 5 min. The cross-linked cells were lysed, and isolated nuclei were digested with 80 U of MboI (NEB, R0543L) at 37 °C for 4 hours. The digested DNA was labeled with biotin-14-dCTP (Invitrogen) and ligated with 4000 U of T4 DNA ligase (NEB) at 20 °C. After cross-link reversal, ligated DNA was extracted using the QIAamp DNA Mini Kit (Qiagen). Purified DNA was sheared to 300–500 bp, blunt-end repaired, A-tailed, and adaptor-ligated. Libraries were then purified through biotin-streptavidin pull-down and PCR amplification. Finally, Hi-C libraries were quantified and sequenced on the Illumina NovaSeq 6000 platform.

Sequencing and data processing. All library construction and sequencing were performed by Berry Genomics Corporation (Beijing, China). Raw sequencing data were processed using standard quality control and trimming protocols before downstream analysis.

Assembly	Total length (Mb)	Number scaffolds/ contigs (chromosomes)	Scaffold/contig N50 length (Mb)	GC (%)	BUSCO (n = 1,367) (%)			
					C	D	F	M
Hifiasm	987.21	644/644	4.10/4.10	32.88	99.9	90.8	0.0	0.1
Purge_Dups	359.25	117/124	5.80/5.80	32.85	99.0	1.5	0.2	0.8
3D-DNA	359.29	484/582 (16)	17.37/4.76	32.85	98.6	1.3	0.1	1.3
Final	288.95	28/111 (16)	18.97/6.25	32.49	98.6	1.3	0.1	1.3

Table 2. Genome assembly statistics for *Thanatophilus sinuatus*. C: complete BUSCOs; D: complete and duplicated BUSCOs; F: fragmented BUSCOs; M: missing BUSCOs.

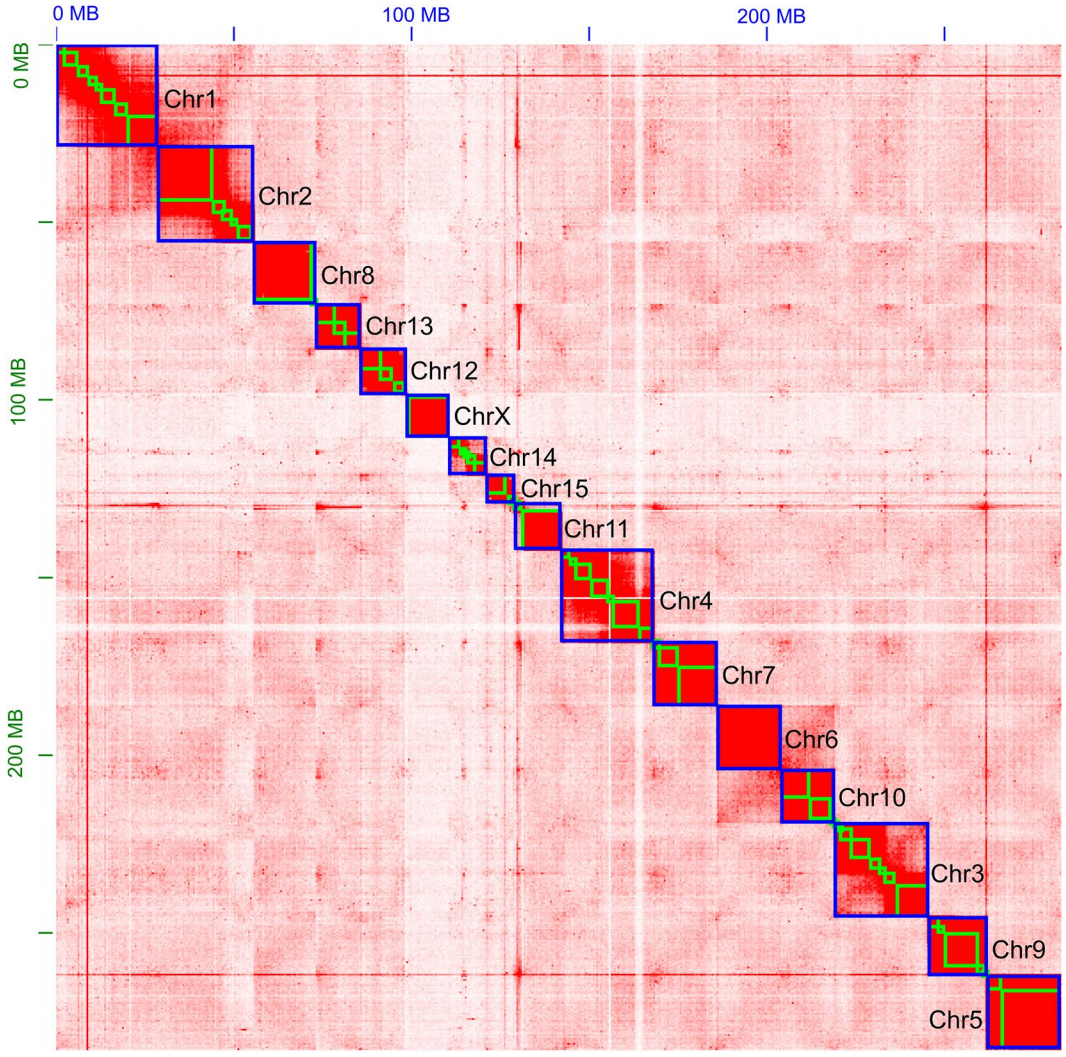

Fig. 2 Genome-wide chromosomal heatmap of *Thanatophilus sinuatus*, with chromosomes in blue and contigs in green.

Sequencing data summary. In total, 104.95 Gb of sequencing data were obtained: 19.74 Gb PacBio HiFi reads (60.31×), 45.70 Gb Hi-C reads (158.16×), 29.83 Gb Illumina reads (103.23 ×), and 9.68 Gb RNA-seq reads. PacBio HiFi reads had a scaffold N50 of 15.94 kb and an average length of 16.02 kb (Table 1).

Genome assembly. Raw Illumina paired-end reads were deduplicated with BBTools v38.82[42] (the "clumpify.sh" script). and stringently trimmed with the "bbduk.sh" script, removing bases with quality scores <20, discarding reads <15 bp, clipping poly-A/G/C tails ≥10 bp, and error-correcting overlapping pairs. We assessed genome size, heterozygosity, and repetitive-sequence content in *T. sinuatus* using a genome-wide k-mer analysis with GenomeScope v2.0[43]. The estimated genome size was 287.23 Mb, displaying a high heterozygosity rate (1.85%) (Fig. 1).

PacBio HiFi data were assembled *de novo* using Hifiasm v0.16.1[44] in "aggressive" mode (with the parameter '-l 2') for removing heterozygous sequences. As low-depth reads frequently represent contamination or erroneous sequences, the Hifiasm assembly retained only contig sequences with a sequencing depth greater than 10X.

Characteristics	*T. sinuatus*
Genome assembly	
Genome Size (Mb)	288.95
Number of scaffolds	28
Number of chromosomes	16
Scaffold N50 length (Mb)	18.97
GC (%)	32.49
BUSCO completeness (%)	98.6
Protein-coding genes	
Number	12,300
Mean gene length (bp)	9,358.8
BUSCO completeness (%)	98.8
Repetitive elements	
Size (Mb)	99.16 (34.32%)
DNA transposons (Mb)	30.38 (10.47%)
SINEs (Mb)	0.30 (0.10%)
LINEs (Mb)	13.36 (4.61%)
LTRs (Mb)	1.26 (0.42%)
Unclassified (Mb)	53.06 (11.70%)
ncRNA	
Number of ncRNA	1755
rRNA	546
miRNA	59
snRNA	116
tRNA	650

Table 3. Genome assembly and annotation statistics of *Thanatophilus sinuatus*.

Residual heterozygous duplication was removed with Purge Dups v1.2.5[45] based on read-depth information from Minimap2 v2.4[46]-mapped BAM files, with a haploid cutoff set at 70 for identifying contigs as haplotigs ("-s 70").

Hi-C reads were aligned to the purged assembly with Juicer v1.6.2[47], and chromosome scaffolding was performed in 3D-DNA v.180922[48]. Assembly breaks were manually inspected and corrected in Juicebox v.1.11.08[47], after which the adjusted assembly was imported into 3D-DNA for final chromosome construction.

Potential contaminants were screened with MMseqs2 v13[49] against the NCBI nucleotide and UniVec databases and with blastn (BLAST + v2.11.020)[50] against UniVec. Sequences showing >90% identity were removed; those with 80–90% identity were verified by online NCBI BLASTN analysis. The potential bacterial and human contaminant scaffolds were excluded.

The resulting *T. sinuatus* assembly spans 288.95 Mb, comprising 111 contigs assembled into 28 scaffolds (GC = 32.49%) (Table 2). Contig and scaffold N50 sizes are 6.25 Mb and 18.97 Mb, respectively. In total, 98.37% of the sequence (284.24 Mb) is anchored to 16 chromosomes (Table 2; Fig. 2).

Genome annotation. A *de novo* repeat library for *T. sinuatus* was built with RepeatModeler v2.0.4[51] (option "-LTRStruct"), which leverages the structural characteristics of repetitive elements and *de novo* prediction methods. The library was merged with RepBase-20181026[52] and Dfam 3.5[53] to create the final repeat set. Repeats were identified and soft-masked with RepeatMasker v.4.1.4[54] against this library, aligning the genome against the final repeat sequence library, revealing that 34.32% of the genome is repetitive: unclassified (9.58%), DNA transposons (10.47%), LINEs (4.61%), simple repeats (8.41%), LTRs (0.42%), and others (Table 3).

Non-coding RNAs (ncRNAs) were detected with Infernal v1.1.4[55] against Rfam v14.10 database[56], and transfer RNAs (tRNAs) with tRNAscan-SE v2.0.9[57]. In total, 1,755 ncRNAs were annotated: 546 rRNAs, 650 tRNAs, 59 microRNAs, 116 small nuclear RNAs and 3 ribozymes (Table 3).

Protein-coding gene annotation was analyzed by MAKER v3.01.03[58] from transcribed RNA, *ab initio* gene predictions, and homologous proteins. Transcribed RNA reads were aligned with HISAT2 v2.2.0[59] and assembled genome-guided with StringTie v2.1.6[60]. *Ab initio* training used BRAKER v2.1.6[61] (GeneMark-ES/ET/EP 4.68_3.60_lic[62] and Augustus v3.3.4[63]) informed by RNA sequence alignments and arthropod reference protein sequences from the OrthoDB10 v1 database[64]. Protein homology and intron position conservation were incorporated via GeMoMa v1.9[65] (parameters "GeMoMa.c = 0.4", "GeMoMa.p = 10") using five reference genomes (*Drosophila melanogaster* (GCF_000001215.4)[66], *Nicrophorus vespilloides* (GCA_001412225.1)[32], *Tribolium castaneum* (GCF_000002335.3)[67], *Apis mellifera* (GCF_003254395.2)[68], and *Bombyx mori* (GCF_014905235.1)[69]. BRAKER and GeMoMa outputs served as ab initio evidence for MAKER. The final annotation comprises 12,300 protein-coding genes (mean length = 9,358.8 bp) containing, on average, 6.2 exons and 5.1 introns (mean exon 314.7 bp; intron 1,515.2 bp; CDS 269.5 bp) (Table 4). The completeness of the protein sequences was evaluated using BUSCO, yielding a high score of 98.8% (n = 1,367). This included

Characteristics	*T. sinuatus*
Structure annotation	
Number of protein-coding genes	12,300
Number of predicted protein sequences	15,562
Mean protein length (aa)	574.6
Mean gene length (bp)	9,358.8
Gene ratio	39.84%
Number of exons per gene	6.2
Mean exon length (bp)	314.7
Exon ratio	8.34%
Number of CDSs per gene	5.9
Mean CDS length (bp)	269.5
CDS ratio	6.88%
Number of introns per gene	5.1
Mean intron length (bp)	1515.2
Intron ratio	31.50%
Function annotation	
Number of genes matching Uniprot records	11,191
Number of genes labeled as "Uncharacterized protein"	193
Number of genes labeled as "unknown function"	1,158
Number of genes with InterProScan annotations	10,138
Number of genes with GO items from InterProScan annotations	6,434
Number of genes with eggNOG annotations	10,914
Number of genes with GO items from eggNOG annotations	8,381
Number of genes with Enzyme Codes (EC) from eggNOG annotations	2,515
Number of genes with KEGG ko terms from eggNOG annotations	7,360
Number of genes with KEGG pathway terms from eggNOG annotations	4,498
Number of genes with COG Functional Categories from eggNOG annotations	10,234
Number of genes with GO items (combining InterProScan and eggNOG results)	9,409
Number of genes with KEGG pathways items (combining InterProScan and eggNOG results)	4,498

Table 4. Genome and functional annotation results of *Thanatophilus sinuatus*.

84.4% (1,154) single-copy, 14.4% (197) duplicated, 0.1% (1) fragmented, and 1.1% (15) missing BUSCOs, reflecting high-quality predictions.

Gene functional annotation was queried against the UniProtKB database using Diamond v2.0.11.149[70] in sensitive mode (parameters "–very-sensitive -e 1e-5"). Gene Ontology (GO) terms, KEGG/Reactome pathways and conserved protein sequences and domains were assigned with eggNOG-mapper v2.1.5[71] and InterProScan 5.53-87.0[72]. The InterProScan analyses included four databases: Pfam[73], SMART[74], Superfamily[75], and CDD[76]. Integrated function predictions yielded 10,234 COG categories, 6,434 GO terms, 7,360 KEGG ko terms, 4,498 KEGG pathways, and 2,515 Enzyme Codes(Table 4). Additionally, chromosome-scale distributions of repeat density, gene density and GC content were plotted with TBtool v1.098769[77] (Fig. 3).

Data Records

Raw sequencing reads and the genome assembly of *Thanatophilus sinuatus* are publicly available at the National Center for Biotechnology Information (NCBI). The Hi-C, transcriptome, Illumina short-read, and PacBio HiFi datasets are archived under the SRA accessions SRR34260723[78], SRR34260724[79], SRR34260725[80], and SRR34260726[81]. The assembled genome is deposited under GenBank accession GCA_051122355.1[82]. The annotation results for repeated sequences, gene structure, and functional prediction have been deposited in figshare[83].

Technical Validation

Assembly quality was appraised with two complementary metrics. First, BUSCO v5.0.4 (Insecta gene set, n = 1,367)[84] recovered 98.6% of expected genes (97.3% single-copy BUSCOs, 1.3% duplicated BUSCOs, 0.1% fragmented BUSCOs, and 1.3% missing BUSCOs). Second, Minimap2/SAMtools read-back mapping showed alignment rates of 98.14% for PacBio HiFi reads and 93.59% for Illumina reads. These assessments collectively validate the high quality of the genome assembly generated in this study.

Data availability

This dataset has not been previously published or presented, and no conflicts of interest are declared. The raw sequencing reads and assembled genome of *Thanatophilus sinuatus* have been deposited in the NCBI under BioProject accession PRJNA1195369. Additionally, annotation files, including repeat elements, gene structures, and functional predictions, are available through Figshare (https://doi.org/10.6084/m9.figshare.29511020).

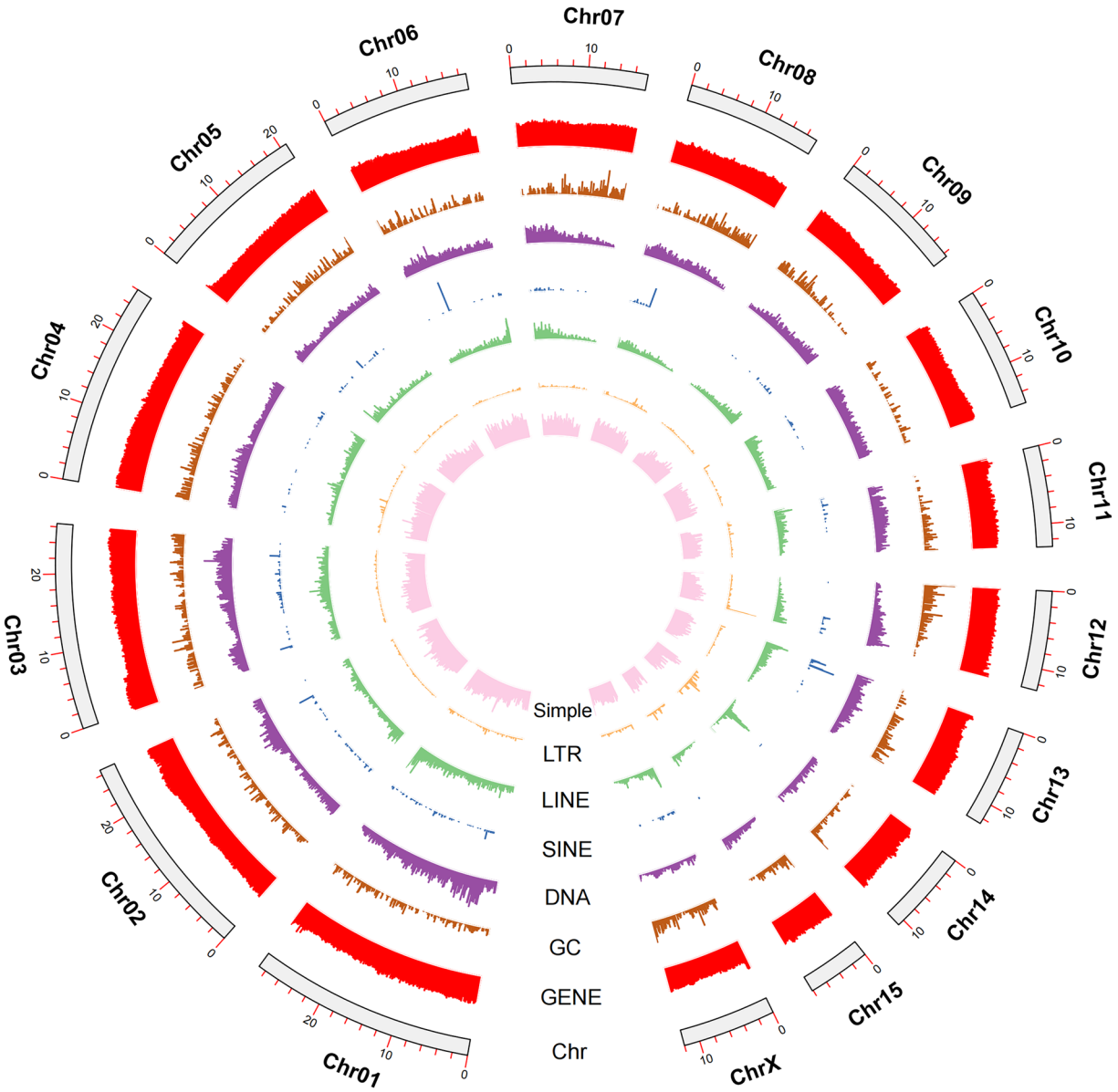

Fig. 3 Overview of *Thanatophilus sinuatus* genome features, showing (from outer to inner rings) chromosome length, gene density, GC content, transposable elements (DNA transposons, short interspersed nuclear elements [SINE], long interspersed nuclear elements [LINE], and long terminal repeats [LTR]), and simple repeats.

Code availability

No specific script was employed in this study. All commands and data-processing pipelines followed the manuals and protocols of the respective bioinformatic software.

Received: 15 July 2025; Accepted: 11 November 2025;
Published online: 10 December 2025

References

1. Kocárek, P. J. B. R. R. Diurnal activity rhythms and niche differentiation in a carrion beetle assemblage (Coleoptera: Silphidae) in Opava, the Czech Republic. **32**, 431-438, https://doi.org/10.1076/brhm.32.4.431.1333 (2001).
2. Qubaiová, J., Jakubec, P., Montoya-Molina, S., Novák, M. & Šuláková, H. The Impact of Diet and Photoperiodism on the Life History of *Thanatophilus sinuatus* (Coleoptera: Silphidae). *Journal of medical entomology* **60**, 453–459, https://doi.org/10.1093/jme/tjad012 (2023).
3. Růžička, J. J. A. S. Z. B. Taxonomic and nomenclatorial notes on *Palaearctic Silphinae* (Coleoptera: Silphidae). **66**, 303–320 (2002).
4. Montoya-Molina, S. *et al*. Developmental Models of the Forensically Important Carrion Beetle, *Thanatophilus sinuatus* (Coleoptera: Silphidae). *Journal of medical entomology* **58**, 1041–1047, https://doi.org/10.1093/jme/tjaa255 (2021).
5. Dekeirsschieter, J., Verheggen, F. J., Haubruge, E. & Brostaux, Y. J. J. O. I. S. Carrion beetles visiting pig carcasses during early spring in urban, forest and agricultural biotopes of Western Europe. **11**, 73, https://doi.org/10.1673/031.011.7301 (2011).
6. Shao, S. *et al*. Research Status of Sarcosaprophagous Beetles as Forensic Indicators. *Insects* **15**, 711, https://doi.org/10.3390/insects15090711 (2024).

7. Jakubec, P., Novák, M., Qubaiová, J., Šuláková, H. & Růžička, J. Description of immature stages of *Thanatophilus sinuatus* (Coleoptera: Silphidae). *International journal of legal medicine* **133**, 1549–1565, https://doi.org/10.1007/s00414-019-02040-1 (2019).
8. Leitch, O., Papanicolaou, A., Lennard, C., Kirkbride, K. P. & Anderson, A. J. B. G. Chemosensory genes identified in the antennal transcriptome of the blowfly *Calliphora stygia*. **16**, 1–17, https://doi.org/10.1186/s12864-015-1466-8 (2015).
9. Rosier, E. *et al*. Time-dependent VOC-profile of decomposed human and animal remains in laboratory environment. *Forensic Science International* **266**, 164–169, https://doi.org/10.1016/j.forsciint.2016.05.035 (2016).
10. Zito, P., Sajeva, M., Raspi, A. & Dötterl, S. J. C. Dimethyl disulfide and dimethyl trisulfide: so similar yet so different in evoking biological responses in saprophilous flies. *Chemoecology* **24**, 261–267, https://doi.org/10.1007/s00049-014-0169-y (2014).
11. Liu, Y. *et al*. Identification of candidate olfactory genes in *Leptinotarsa decemlineata* by antennal transcriptome analysis. *Frontiers in Ecology and Evolution* **3**, 60, https://doi.org/10.3389/fevo.2015.00060 (2015).
12. Ren, L. *et al*. Chromosome-level genome assembly of the synanthropic fly *Chrysomya megacephala*: insights into oviposition location. *BMC Genomics* **26**, 442, https://doi.org/10.1186/s12864-025-11645-3 (2025).
13. Xu, L. *et al*. Two odorant receptors regulate 1-octen-3-ol induced oviposition behavior in the oriental fruit fly. **6**, 176 (2023).
14. Jürgens, A., Wee, S. L., Shuttleworth, A. & Johnson, S. D. J. E. L. Chemical mimicry of insect oviposition sites: a global analysis of convergence in angiosperms. *Ecology Letters* **16**, 1157–1167, https://doi.org/10.1111/ele.12152 (2013).
15. Dekeirsschieter, J. *et al*. Electrophysiological and behavioral responses of *Thanatophilus sinuatus* Fabricius (Coleoptera: Silphidae) to selected cadaveric volatile organic compounds. *Journal of Forensic Sciences* **58**, 917–923 (2013).
16. Frątczak, K. & Matuszewski, S. Classification of forensically-relevant larvae according to instar in a closely related species of carrion beetles (Coleoptera: Silphidae: Silphinae). *Forensic science, medicine, and pathology* **12**, 193–197, https://doi.org/10.1007/s12024-016-9774-0 (2016).
17. Shao, S. *et al*. Developmental time pattern of *Thanatophilus sinuatus* at different constant and variable temperatures. *Forensic science international* **366**, 112301, https://doi.org/10.1016/j.forsciint.2024.112301 (2025).
18. Lis, N., Szyk, W., Mądra-Bielewicz, A. & Matuszewski, S. Calibrating insect age at eclosion by size in a gregarious carrion beetle *Thanatophilus sinuatus* (Staphylinidae: Silphinae). *Medical and veterinary entomology* **37**, 705–714, https://doi.org/10.1111/mve.12674 (2023).
19. Grebennikov, V., Newton, A. J. A. S. & Phylogeny. Detecting the basal dichotomies in the monophylum of carrion and rove beetles (Insecta: Coleoptera: Silphidae and Staphylinidae) with emphasis on the Oxyteline group of subfamilies. *Arthropod Systematics & Phylogeny* **70**, 133–165 (2012).
20. Sikes, D. S., Thayer, M. K. & Newton, A. F. J. Z. Large carrion and burying beetles evolved from Staphylinidae (Coleoptera, Staphylinidae, Silphinae): a review of the evidence. *Zookeys* **1200**, 159, https://doi.org/10.3897/zookeys.1200.122835 (2024).
21. Mckenna, D. D. *et al*. Phylogeny and evolution of *S taphyliniformia* and *S carabaeiformia*: forest litter as a stepping stone for diversification of nonphytophagous beetles. *Systematic Entomology* **40**, 35–60, https://doi.org/10.1111/syen.12093 (2015).
22. Mckenna, D. D. *et al*. The beetle tree of life reveals that *C oleoptera* survived end-P ermian mass extinction to diversify during the C retaceous terrestrial revolution. *Systematic Entomology* **40**, 835–880, https://doi.org/10.1111/syen.12132 (2015).
23. Zhang, S.-Q. *et al*. Evolutionary history of Coleoptera revealed by extensive sampling of genes and species. *Nature Communications* **9**, 205, https://doi.org/10.1038/s41467-017-02644-4 (2018).
24. Zhao, T.-Y., He, L., Xu, X., Chen, Z.-N. & Gao, Y.-Y. J. Z. The first mitochondrial genome of *Creophilus Leach* and *Platydracus Thomson* (Coleoptera: Staphylinidae: Staphylinini) and phylogenetic implications. *Zootaxa* **5099**, 179–200, https://doi.org/10.11646/zootaxa.5099.2.2 (2022).
25. Růžička, J., Jakubec, P., Mahlerová, K., Šípková, H. & Nishikawa, M. J. S. R. Integrative taxonomy and species distribution models of the genus Diamesus Hope, 1840 (Coleoptera: Staphylinidae: Silphinae). *Scientific Reports* **13**, 3192, https://doi.org/10.1038/s41598-023-30019-x (2023).
26. Cai, C. *et al*. Integrated phylogenomics and fossil data illuminate the evolution of beetles. *Royal Society Open Science* **9**, 211771, https://doi.org/10.1098/rsos.211771 (2022).
27. Mahlerová, K., Jakubec, P., Krak, K. & Růžička, J. J. S. E. Resolving the intergeneric phylogeny of the large carrion beetles (Staphylinidae: Silphinae: Silphini). *Systematic Entomology* **50**, 168–179, https://doi.org/10.1111/syen.12650 (2025).
28. Kadlec, J., Jakubec, P. & Šípek, P. J. J. O. M. E. *In vitro* comparison of two food choice methodologies for necrophagous species of the genus Thanatophilus (Coleoptera: Staphylinidae: Silphinae). *Journal of Medical Entomology* **61**, 603–610, https://doi.org/10.1093/jme/tjae018 (2024).
29. Büchner, G. *et al*. Ecological drivers of carrion beetle (Staphylinidae: Silphinae) diversity on small to large mammals. *Ecology and Evolution* **14**, e70203, https://doi.org/10.1002/ece3.70203 (2024).
30. Kim, M. I., Park, J. S., Choung, C. M., Kim, M. J. & Kim, I. J. J. O. A.-P. E. Complete mitochondrial genome of the forensically important carrion beetle, *Necrodes nigricornis* (Coleoptera: silphidae). *Journal of Asia-Pacific Entomology* **26**, 102033, https://doi.org/10.1016/j.aspen.2022.102033 (2023).
31. Jiang, Y. *et al*. The complete mitochondrial genome of a corpse related necrophagous beetle, *Necrodes littoralis* (Coleoptera: Silphidae). *Mitochondrial DNA B Resour.* **6**, 1512–1513, https://doi.org/10.1080/23802359.2021.1914236 (2021).
32. Cunningham, C. B. *et al*. The genome and methylome of a beetle with complex social behavior, *Nicrophorus vespilloides* (Coleoptera: Silphidae). *Genome Biology and Evolution* **7**, 3383–3396, https://doi.org/10.1093/gbe/evv194 (2015).
33. Parker, D. J. *et al*. Transcriptomes of parents identify parenting strategies and sexual conflict in a subsocial beetle. *Nature Communications* **6**, 8449, https://doi.org/10.1038/ncomms11737 (2015).
34. Crowley, L. M. *et al*. The genome sequence of the Banded Burying beetle, *Nicrophorus investigator* Zetterstedt, 1824. *Wellcome Open Research* **9**, 343, https://doi.org/10.12688/wellcomeopenres.21496.1 (2024).
35. Herndon, N. *et al*. Enhanced genome assembly and a new official gene set for *Tribolium castaneum*. *BMC Genomics* **21**, 1–13, https://doi.org/10.1186/s12864-019-6394-6 (2020).
36. Vavricka, C. J. *et al*. Tyrosine metabolic enzymes from insects and mammals: a comparative perspective. *Insect Science* **21**, 13–19, https://doi.org/10.1111/1744-7917.12038 (2014).
37. Jin, Y. *et al*. Differential gene expression analysis of the *Epacromius coerulipes* (Orthoptera: Acrididae) transcriptome. *Journal of Insect Science* **16**, 42, https://doi.org/10.1093/jisesa/iew014 (2016).
38. Liang, H. & Song, K. J. P. O. Elucidating ascorbate and aldarate metabolism pathway characteristics via integration of untargeted metabolomics and transcriptomics of the kidney of high-fat diet-fed obese mice. *PLoS One* **19**, e0300705, https://doi.org/10.1371/journal.pone.0300705 (2024).
39. Mahlerová, K., Jakubec, P., Novák, M. & Růžička, J. J. S. R. Description of larval morphology and phylogenetic relationships of *Heterotemna tenuicornis* (Silphidae). *Scientific Reports* **11**, 16973, https://doi.org/10.1038/s41598-021-94744-x (2021).
40. Bonacci, T. *et al*. Insect fauna and degradation activity of *Thanatophilus* species on carrion in southern Italy (Coleoptera: Silphidae). *Entomologia Generalis* **33**, 63–70 (2011).
41. 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).
42. Bushnell, B. BBtools. Available online: https://sourceforge.net/projects/bbmap/ (accessed on 1 October 2022) (2014).
43. Ranallo-Benavidez, T. R., Jaron, K. S. & Schatz, M. C. J. N. 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).

44. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. J. N. m. 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).
45. 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).
46. Li, H. J. B. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100, https://doi.org/10.1093/bioinformatics/bty191 (2018).
47. 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).
48. 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).
49. Steinegger, M. & Söding, J. J. N. B. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* **35**, 1026–1028, https://doi.org/10.1038/nbt.3988 (2017).
50. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. J. J. O. M. B. Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403–410, https://doi.org/10.1016/S0022-2836(05)80360-2 (1990).
51. Flynn, J. M. *et al*. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci USA* **117**, 9451–9457, https://doi.org/10.1073/pnas.1921046117 (2020).
52. Bao, W., Kojima, K. K. & Kohany, O. J. M. D. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob DNA* **6**, 1–6, https://doi.org/10.1186/s13100-015-0041-9 (2015).
53. Hubley, R. *et al*. The Dfam database of repetitive DNA families. *Nucleic Acids Res* **44**, D81–D89, https://doi.org/10.1093/nar/gkv1272 (2016).
54. Smit *et al*. 2013–201526.Smit, A. F. A., Hubley, R. & Green, P. RepeatMasker Open-4.0. Available online: http://www.repeatmasker.org (accessed on 1 October 2022) (2013–2015).
55. Nawrocki, E. P. & Eddy, S. R. J. B. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29**, 2933–2935, https://doi.org/10.1093/bioinformatics/btt509 (2013).
56. Griffiths-Jones, S. *et al*. Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res* **33**, D121–D124, https://doi.org/10.1093/nar/gki081 (2005).
57. Chan, P. P. & Lowe, T. M. tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. *Methods in Molecular Biology* **1962**, 1–14, https://doi.org/10.1007/978-1-4939-9173-0_1 (2019).
58. Holt, C. & Yandell, M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC bioinformatics* **12**, 491, https://doi.org/10.1186/1471-2105-12-491 (2011).
59. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nature methods* **12**, 357–360, https://doi.org/10.1038/nmeth.3317 (2015).
60. 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).
61. 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 genomics and bioinformatics* **3**, lqaa108, https://doi.org/10.1093/nargab/lqaa108 (2021).
62. Brůna, T., Lomsadze, A. & Borodovsky, M. GeneMark-EP+: eukaryotic gene prediction with self-training in the space of genes and proteins. *NAR genomics and bioinformatics* **2**, lqaa026, https://doi.org/10.1093/nargab/lqaa026 (2020).
63. Stanke, M., Steinkamp, R., Waack, S. & Morgenstern, B. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic acids research* **32**, W309–312, https://doi.org/10.1093/nar/gkh379 (2004).
64. 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 research* **47**, D807–d811, https://doi.org/10.1093/nar/gky1053 (2019).
65. Keilwagen, J., Hartung, F., Paulini, M., Twardziok, S. O. & Grau, J. Combining RNA-seq data and homology-based gene prediction for plants, animals and fungi. *BMC bioinformatics* **19**, 189, https://doi.org/10.1186/s12859-018-2203-5 (2018).
66. Hoskins, R. A. *et al*. The Release 6 reference sequence of the *Drosophila melanogaster* genome. *Genome research* **25**, 445–458, https://doi.org/10.1101/gr.185579.114 (2015).
67. Richards, S. *et al*. The genome of the model beetle and pest *Tribolium castaneum*. *Nature* **452**, 949–955, https://doi.org/10.1038/nature06784 (2008).
68. Insights into social insects from the genome of the honeybee Apis mellifera. *Nature* **443**, 931–949, https://doi.org/10.1038/nature05260 (2006).
69. Kawamoto, M. *et al*. High-quality genome assembly of the silkworm, Bombyx mori. *Insect biochemistry and molecular biology* **107**, 53–62, https://doi.org/10.1016/j.ibmb.2019.02.002 (2019).
70. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nature methods* **12**, 59–60, https://doi.org/10.1038/nmeth.3176 (2015).
71. 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).
72. Finn, R. D. *et al*. InterPro in 2017-beyond protein family and domain annotations. *Nucleic acids research* **45**, D190–d199, https://doi.org/10.1093/nar/gkw1107 (2017).
73. El-Gebali, S. *et al*. The Pfam protein families database in 2019. *Nucleic acids research* **47**, D427–d432, https://doi.org/10.1093/nar/gky995 (2019).
74. Letunic, I. & Bork, P. 20 years of the SMART protein domain annotation resource. *Nucleic acids research* **46**, D493–d496, https://doi.org/10.1093/nar/gkx922 (2018).
75. Wilson, D. *et al*. SUPERFAMILY–sophisticated comparative genomics, data mining, visualization and phylogeny. *Nucleic acids research* **37**, D380–386, https://doi.org/10.1093/nar/gkn762 (2009).
76. Marchler-Bauer, A. *et al*. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic acids research* **45**, D200–d203, https://doi.org/10.1093/nar/gkw1129 (2017).
77. Chen, C. *et al*. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Molecular plant* **13**, 1194–1202, https://doi.org/10.1016/j.molp.2020.06.009 (2020).
78. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR34260723 (2025).
79. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR34260724 (2025).
80. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR34260725 (2025).
81. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRR34260726 (2025).
82. *NCBI Assembly* https://identifiers.org/ncbi/insdc.gca:GCA_051122355.1 (2025).
83. Zhang, Y. Genome annotation of of *Thanatophilus sinuatus*. *figshare. Dataset*. https://doi.org/10.6084/m9.figshare.29511020 (2025).
84. Waterhouse, R. M. *et al*. BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. *Molecular biology and evolution* **35**, 543–548, https://doi.org/10.1093/molbev/msx319 (2018).

Acknowledgements
This study was supported by grants from the National Natural Science Foundation of China (grant numbers 82160323, 32270545, 32070508, 82002007, 82170542, and 92251307), the Major project of the National Natural Science Foundation of China (grant numbers 82293650 and 82293651), and the Open project of Hebei Key Laboratory of Forensic Medicine (grant number JYFY-23KF006). We gratefully acknowledge Mr. Hongwei Liu at Berry Genomics Corporation (Beijing, China) for providing technical assistance in this research.

Author contributions
R.Z. and Y.W. contributed to the research design. S.S. and L.L. collected the samples. J.J. and Y.Z. analyzed the data. Y.Z., Y.W. and J.J. 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.

Additional information
Correspondence and requests for materials should be addressed to Y.W. or R.Z.

Reprints and permissions information is available at www.nature.com/reprints.

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