scientific data

Check for updates

OPEN

DATA DESCRIPTOR

The first chromosome-level genome of the stag beetle *Dorcus hopei* Saunders, 1854 (Coleoptera: Lucanidae)

Xiaolu Li[1,4], Chuyang Mao[2,3,4], Jinwu He[2,3], Xiaoyan Bin[1], Guichun Liu[2,3], Zhiwei Dong[2,3], Ruoping Zhao[2,3], Xia Wan[1]✉ & Xueyan Li[2,3]✉

Stag beetles (Coleoptera: Lucanidae) represent a significant saproxylic assemblage in forest ecosystems and are noted for their enlarged mandibles and male polymorphism. Despite their relevance as ideal models for the study of exaggerated mandibles that aid in attracting mates, the regulatory mechanisms associated with these traits remain understudied, and restricted by the lack of high-quality reference genomes for stag beetles. To address this limitation, we successfully assembled the first chromosome-level genome of a representative species *Dorcus hopei*. The genome was 496.58 Mb in length, with a scaffold N50 size of 54.61 Mb, BUSCO values of 99.8%, and 96.8% of scaffolds anchored to nine pairs of chromosomes. We identified 285.27 Mb (57.45%) of repeat sequences and annotated 11,231 protein-coding genes. This genome will be a valuable resource for further understanding the evolution and ecology of stag beetles, and provides a basis for studying the mechanisms of exaggerated mandibles through comparative analysis.

Background & Summary

Stag beetles (Family: Lucanidae), comprise over 1,800 species and subspecies[1], noted for enlarged allometry mandibles and male polymorphism[2]. As a holometabolous insect, they undergo a complete metamorphosis with four life stages: egg, larva, pupa, and adult[3,4]. The larvae primarily feed on decaying wood, while the adults are mostly nocturnal and feed on plant juices, fruits, or other decaying organic matter[5–7]. Due to their saprophagous nature, stag beetles play essential roles in the carbon and nitrogen cycles, which also makes stag beetles an important indicator species for evaluating forest ecosystems[1,8,9]. Based on observations of male lucanid beetles, Darwin (1871)[10] noted that "The great mandibles of the male Lucanidae are extremely variable both in size and structure…and are used as efficient weapons for fighting". It implies that individuals with larger mandibles have a better chance of defeating their rivals and winning mating rights[11,12]. Due to its unique and variable appearances, as well as interesting behavioral phenomena, this group has garnered the affection of many collectors, entomologists and evolutionary biologists[8,10].

Stag beetles are widely distributed across various biogeographic regions and represent a group with significant nodal importance in the process of evolution[13]. The research on the Lucanidae family is predominantly concentrated on molecular taxonomic studies[8,14,15], based on the data from nuclear gene fragments[16], mitochondrial multi-gene fragments[17] and mitochondrial genomes[18,19]. However, these data are insufficient to provide more insights into the formation and differentiation of the stag beetles' mandibles[20]. Decoding high-quality reference genomes has been proven to be the cornerstone of inferring phylogeny and exploring the molecular basis behind phenotypic innovation[21], e.g., the antlers of cervids[22], the long tail feathers of birds[23,24], and the horns of some scarabs[25,26]. The limited availability of genomic data hindered our research on Lucanidae family.

[1]Department of Ecology, School of Resources and Environmental Engineering, Anhui University, Hefei, 230601, China. [2]Key Laboratory of Genetic Evolution & Animal Models, Kunming Institute of Zoology, Chinese Academy of Sciences (CAS), Kunming, Yunnan, 650223, China. [3]Yunnan Key Laboratory of Biodiversity Information, Kunming, Yunnan, 650223, China. [4]These authors contributed equally: Xiaolu Li, Chuyang Mao. ✉e-mail: wanxia@ahu.edu.cn; lixy@mail.kiz.ac.cn

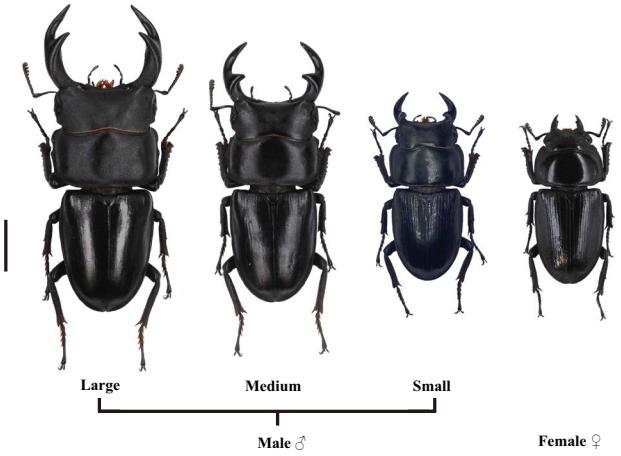

Fig. 1 Sexual dimorphism and male trimorphism in *Dorcus hopei*. The scale bar is 1 cm.

Dorcus hopei (Saunders, 1854), distributing from central and northeastern China, is a well-known species notable for its sword-shaped mandibles[2,27,28] (Fig. 1). Comparatively, its mandibles are simpler to observe, with a large sharp bump and a relatively small inner tooth. Owing to the restriction of insect allometry and scaling relationship[29–31], its male trimorphism in mandibles and body sizes is a very rare type. Based on these characteristics, *D. hopei* is a good choice for performing long-term studies in stag beetles.

In this study, we successfully assembled the first chromosome-level reference genome of *D. hopei* using Illumina, Nanopore and Hi-C sequencing, the information from which could enhance our understanding of stag beetle survival and evolution. Furthermore, it provides a novel clue for uncovering the molecular basis of extreme mandibles development and male trimorphism formation in the future.

Methods

Sample information. Adult male *D. hopei* specimens were collected from Shou County, Huainan City, Anhui Province, China, during May and June from 2017 to 2021. The beetles were subsequently reared in the laboratory (23 °C, 14 h:10 h light/dark cycle, and 45% relative humidity) and provided with brown sugar jelly and bananas as food. Two adult males were selected for next-generation genomic sequencing using the Illumina platform, one adult male was used for long-read genomic sequencing with the Oxford Nanopore platform, two adult males were selected for Hi-C sequencing, and one adult male was used for transcriptomic sequencing.

Illumina, nanopore, Hi-C, and RNA sequencing. Genomic DNA was isolated from the leg muscles using a Trelief Animal Genomic DNA Kit (TsingKe, China). Paired-end libraries (insert size: 350 bp) were generated using a NEBNext Ultra DNA Library Prep Kit (New England Biolabs, USA) with the Illumina HiSeq 4000 platform at Novogene (Tianjin, China). After filtering the bases in the raw reads of quality $<$Q20, we obtained 55.13 Gb (113x) clean Illumina data.

For Oxford Nanopore long-read sequencing, DNA from thorax muscles were extracted using a Qiagen DNAeasy Kit (Qiagen, German). Subsequently, the extracted DNA was treated with the NEBNext Ultra End Repair/dA-Tailing module (New England Biolabs, USA) to incorporate adapters for priming sequencing reactions (NextOmics, China). The library was constructed using a 1D DNA Ligation Sequencing Kit (SQK-LSK109) (Oxford Nanopore Technologies, England) and sequencing was performed on a PromethION flow cell (NextOmics, China) to obtain 46.34 Gb (95x) Nanopore data.

For Hi-C sequencing, cells isolated from head tissues were fixed with formaldehyde and subsequently digested using the restriction enzyme MboI. The DNA was purified and then sheared into 300–600 bp fragments using a Covaris M220 device (Covaris, USA). After DNA size selection using AMPure XP beads, point ligation junctions were pulled down using Dynabeads MyOne Streptavidin C1 (ThermoFisher, USA). Then the Hi-C library was sequenced on the Illumina NovaSeq sequencing platform at Novogene (China), and we got 48.46 Gb (100x) Hi-C data.

Transcriptomic sequencing was used to assist in gene structure annotation. RNA was extracted from the head tissue of one adult male using TRIzol. RNA quality was assessed using an RNA Nano 6000 Assay Kit for 2100 Bioanalyzer Systems Kit (Agilent Technologies, China). The libraries were generated using a NEBNext Ultra RNA Library Prep Kit (New England Biolabs, USA) and sequenced on the Illumina Hiseq platform at Novogene (Tianjin, China). And 8.07 Gb were obtained for assisting genomic annotation.

Chromosome-level genome assembly. Illumina data was used to estimate genome size based on 17 *k-mer* size analysis using KmerFreq v5.0[32]. The estimated genome size of *D. hopei* was 487.15 Mb, with heterozygosity of 0.021 based on the frequency distribution of 17-mers (Fig. 2a). The Oxford Nanopore long reads were used to assemble and polish the primary genome with NextDenovo v2.5.0 (https://github.com/Nextomics/NextDenovo) (parameters: -k 0 -p 15) and purge_dups v1.0.0[33]. The Illumina short reads were used to correct errors at the base level in the above-polished genome using NextPolish v1.4.0[34]. The Nanopore assembly was

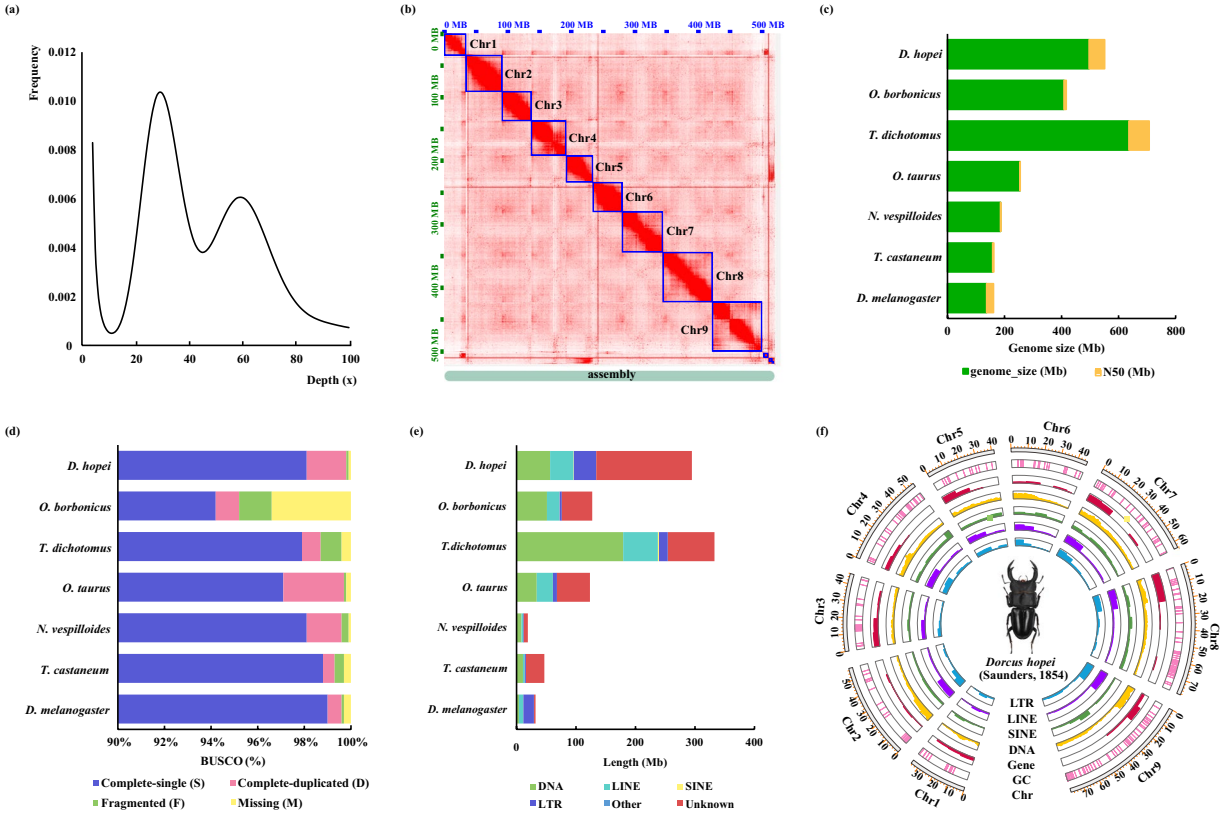

Fig. 2 Assembly of chromosome-level genome of *Dorcus hopei*. (**a**) 17-mer analysis of the *D. hopei* genome based on Illumina reads, X-axis represented depth (x); Y-axis represented the proportion of the frequency of that depth to the total frequency of all depths. (**b**) Heatmap of Hi-C data showing nine chromosome boundaries (Chr1 to Chr9). The comparison of (**c**) genome size and N50 length, (**d**) BUSCO scores, and (**e**) repeat elements in *D. hopei* and six other species. (**f**) Circos tracks showing chromosome length, GC content, density of protein-coding genes, and repetitive elements (SINE, short interspersed elements; LINE, long interspersed elements; LTR, long terminal repeat elements).

496.47 Mb (N50 = 3.94 Mb), comprised 232 contigs, and achieved a BUSCO completeness score of 99.90% (Table 1).

The Hi-C paired-end reads were iteratively mapped to the Nanopore assembly using HiC-Pro v2.9.0[35]. The paired tags were then filtered using restriction enzyme digesting fragments with Juicer v1.60 and contigs were ordered and orientated using 3D *de novo* assembly software (3D-DNA) v180922[36]. Finally, JuiceBox v1.11.08[37] was applied to correct contig orientation and move suspicious fragments into unanchored groups by visual exploration of the Hi-C heatmap. After Hi-C assembly, the resulting 496.58 Mb genome was assembled into 18 chromosomes (2n = 8AA + XY) (Table 1; Fig. 2b). Notably, 96.18% of the contigs from the "Nanopore assembly" were successfully anchored to nine chromosomes, with a scaffold N50 of 54.61 Mb and 99.80% BUSCO completeness (1.7% duplicated genes) (Table 1), indicating relatively high assembly integrity (Fig. 2c,d).

Genome annotation. We choose several reference species to assist annotation, including five other coleopteran species (Scarabaeoidea: *Onthophagus taurus* (GCA_000648695.2), *Oryctes borbonicus*[38] (GCA_902654985.2), *Trypoxylus dichotomus*[39] (GCA_023509865.1); Staphylinidea: *Nicrophorus vespilloides*[40] (GCA_001412225.1); Tenebrionoidea: *Tribolium castaneum*[41] (GCA_000002335.3)), and one dipteran species, *Drosophila melanogaster*[42] (GCA_000001215.4). We uploaded the detailed species information table to figshare[43]. Initially, we annotated repetitive sequences in the *D. hopei* genome by identifying LTRs and tandem repeats using LTR_Finder v1.05[44] and Tandem Repeat Finder v4.07b[45], respectively. Transposable elements (TEs), including DNA elements, long interspersed nuclear elements (LINEs), short interspersed nuclear elements (SINEs), and long terminal repeats (LTRs), were next identified using RepeatMasker v4.0.5[46] against a *de novo* repeat library constructed with RepeatModeler v1.0.4[47] and Repbase TE library v16.02[48] separately at the DNA level. Finally, TE-relevant proteins were identified using RepeatProteinMask v4.0.9[47] at the protein level. The final genome assembly (Hi-C assembly) of *D. hopei* comprised 57.45% repetitive sequences, totaling approximately 285.27 Mb, which is almost twice that of *T. castaneum* (31.15%) (Fig. 2e). Among the repetitive sequences in the *D. hopei* genome, the major categories included unclassified sequences (32.39%), DNA elements (11.36%) with maximum density in each chromosome, LINEs (7.95%), and LTRs (7.55%) (Fig. 2f).

Protein-coding genes were predicted using a combination of *de novo*-, homology-, and transcriptome-based approaches. We utilized the repeat-masked genome and applied the *de novo*-based gene prediction software

Statistics	Nanopore assembly	Hi-C assembly
Assembly		
Total number	232	27
Genome size (Mb)	496.47	496.58
Average length (Mb)	2.14	18.39
N50 length (Mb)	3.94	54.61
N90 length (Mb)	1.01	33.12
Maximum length (Mb)	17.85	75.00
GC content (%)	35.23	35.22
BUSCO (completeness, %)	99.9	99.8
Anchor rate (%)	—	96.18
Illumina reads mapping rate (properly paired) (%)	99.10 (94.20)	99.10 (94.20)
Nanopore reads mapping rate (%)	88.81	100.00
Annotation		
Gene number		11,231
Average gene length (bp)		15,840.00
Average CDS[a] length (bp)		1,714.51
Average exon number		6.58
Average exon length (bp)		260.59
Average intron length (bp)		2,531.69
BUSCO (%)		92.0

Table 1. Statistics for *Dorcus hopei* genome assembly and gene annotation. [a]CDS: coding sequence.

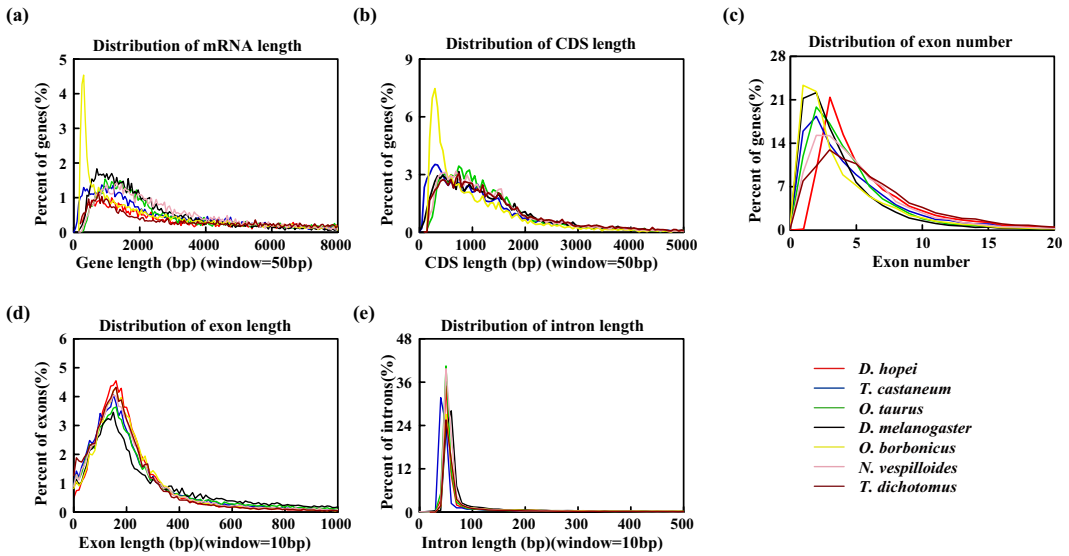

Fig. 3 Distribution statistics of gene features among the seven species. The comparison of (**a**) mRNA length, (**b**) CDS length, (**c**) exon number, (**d**) exon length and (**e**) intron length in *D. hopei* and other six species.

Augustus v3.4.0[49], using models trained on protein sequences from the *O. borbonicus* genome[38], with default parameters. TBLASTN v2.12.0[50] and GeneWise v2.4.1[51] were used for homology prediction. The transcriptome data were then aligned to the genome using HISAT2 v2.0.0-beta[52]. Based on the resulting BAM files and reference genome, the transcriptomic sequences were assembled using StringTie v2.1.4[53]. To form a comprehensive, non-redundant set of genes, we performed several integrations using EVidenceModeler (EVM) v1.1.1[54], assigning different weight values to the seven genomes based on their BUSCO scores and gene structure components (gene length, coding sequence length, exon number and length, and intron length). The EVM gene set with the best BUSCO value and gene structure components was then selected as the final gene prediction. Finally, resulting in the annotation of 11,231 protein-coding genes in the *D. hopei* genome. We uploaded the complete gene annotation tables to figshare[43]. Compared to the different gene features of other six species, the *D. hopei* genome annotations were comprehensive (Fig. 3), further validating the quality and accuracy of the genome annotation.

Finally, we performed functional annotation of the genome. The protein sequences of the genome were searched for homology-based function assignments against the KEGG, NR, TrEMBL, and SwissProt databases using BLASTP v2.2.26[55] with an e-value cut-off of 1e-5. Domains in the *D. hopei* genome using InterProScan

Functional database	Number of genes annotated (11,231)
Annotated	9,942 (88.52%)
Unannotated	1,289 (11.48%)
InterPro	8,737 (77.79%)
GO	8,737 (77.79%)
KEGG	6,556 (58.37%)
SwissProt	7,654 (68.15%)
TrEMBL	9,779 (87.07%)
NR	9,838 (87.60%)

Table 2. Statistics of functional annotation of the *Dorcus hopei* protein-coding genes.

v5.54–87.0[56] with InterPro and GO database. And combined above results, 88.52% of the predicted genes were functionally annotated using six functional protein databases (Table 2).

Data Records

The chromosome-level assembly and annotation file of *D. hopei* has been deposited in figshare database[57]. Raw sequencing data (Illumina reads, Nanopore reads, Hi-C reads, RNA-seq reads) and sample information are available at NCBI, which can be found under identification number SRP440764[58]. The assembly also has been deposited in NCBI with the accession number GCA_033060865.1[59]. More detailed information about selected species, the results of genomic annotation (repeated sequences and gene structure), orthologs, and synteny has been deposited in figshare database[43].

Technical Validation

Quality assessment of the assembled genome was performed using the following methods. Firstly, BWA v0.7.17[60] was used to map the Illumina reads to the *D. hopei* assembly and Samtools v1.3.1[61] was used to calculate the mapping ratio. The Illumina short reads with a 99.10% accuracy ratio were mapped to the final assembly (Table 1). Secondly, compared N50 length/number with other six selected species. The *D. hopei* genome displayed a longer N50 (54.61 Mb) and better continuity compared to the chromosome-level genomes of *T. castaneum* and *T. dichotomus* (Fig. 2c). Thirdly, insecta_odb10 with 1,367 genes in BUSCO v5.2.2[62] was used to evaluate genome assembly and annotation completeness. The final assembly had 99.8% BUSCO scores with 0.1% fragmented and 0.1% missing sequences (Fig. 2d). Additionally, we got nine pairs of chromosomes based on Hi-C data, mirroring that of congeneric species *Dorcus parallelipipedus*[63]. All these results suggest that we got a high-quality assembly of *D. hopei* with high integrity, continuity and accuracy.

Code availability

No custom code was used in this study. The data analyses used standard bioinformatic tools specified in the methods.

Received: 5 October 2023; Accepted: 10 April 2024;
Published online: 18 April 2024

References

1. Wan, X., Jiang, Y., Cao, Y., Sun, B. & Xiang, X. Divergence in gut bacterial community structure between male and female stag beetles *Odontolabis fallaciosa* (Coleoptera, Lucanidae). *Animals.* **10**, 2352 (2020).
2. Arai, T. *et al.* Discovery of hyperactive antifreeze protein from phylogenetically distant beetles questions its evolutionary origin. *Int J Mol Sci.* **22**, 3637 (2021).
3. Goyens, J., Dirckx, J. & Aerts, P. Jaw morphology and fighting forces in stag beetles. *J Exp Biol.* **219**, 2955–2961 (2016).
4. Gotoh, H. *et al.* Juvenile hormone regulates extreme mandible growth in male stag beetles. *PLoS One.* **6**, e21139 (2011).
5. Harvey, D. J., Gange, A. C., Hawes, C. J. & Rink, M. Bionomics and distribution of the stag beetle, *Lucanus cervus* (L.) across Europe*. *Insect Conserv Divers.* **4**, 23–38 (2011).
6. Kubota, K. *et al.* Evolutionary relationship between *Platycerus* stag beetles and their mycangium-associated yeast symbionts. *Front Microbiol* **11**, 1436 (2020).
7. Hendriks, P. Life cycle length of the lesser stag beetle (Coleoptera: Lucanidae: *Dorcus parallelipipedus*). *Entomol Ber.* **79**, 208–216 (2019).
8. Kim, S. I. & Farrell, B. D. Phylogeny of world stag beetles (Coleoptera: Lucanidae) reveals a Gondwanan origin of Darwin's stag beetle. *Mol Phylogenet Evol.* **86**, 35–48 (2015).
9. Tanahashi, M., Ikeda, H. & Kubota, K. Elementary budget of stag beetle larvae associated with selective utilization of nitrogen in decaying wood. *Sci Nat.* **105**, 33 (2018).
10. Darwin, C. *The Descent Of Man, And Selection In Relation To Sex, Vol 1* (John Murray, 1871).
11. Pennisi, E. Insulin may guarantee the honesty of beetle's massive horn. *Science.* **337**, 408 (2012).
12. Mills, M. R. *et al.* Functional mechanics of beetle mandibles: honest signaling in a sexually selected system. *J Exp Zool A Ecol Integr Physiol.* **325A**, 3–12 (2016).
13. Zhang, M. *et al.* Geometric morphometric analysis of the pronotum and elytron in stag beetles: insight into its diversity and evolution. *Zookeys.* **833**, 21–40 (2019).
14. Bartolozzi, L., Norbiato, M. & Cianferoni, F. A review of geographical distribution of the stag beetles in Mediterranean countries (Coleoptera: Lucanidae). *Fragm Entomol.* **48**, 153–168 (2016).
15. Méndez, M. & Thomaes, A. Biology and conservation of the European stag beetle: recent advances and lessons learned. *Insect Conserv Divers.* **14**, 271–284 (2021).

16. Kubota, K. *et al*. Diversification process of stag beetles belonging to the genus *Platycerus* Geoffroy (Coleoptera: Lucanidae) in Japan based on nuclear and mitochondrial genes. *Entomol Sci.* **14**, 411–427 (2011).
17. Yuan, J. J., Chen, D. & Wan, X. A multilocus assessment reveals two new synonymies for East Asian *Cyclommatus* stag beetles (Coleoptera, Lucanidae). *ZooKeys* **1021**, 65–79 (2021).
18. Lin, Z.-Q., Song, F., Li, T., Wu, Y.-Y. & Wan, X. New mitogenomes of two Chinese stag beetles (Coleoptera, Lucanidae) and their implications for systematics. *J Insect Sci.* **17**, 63 (2017).
19. Zeng, L. *et al*. Comparative mitochondrial genomics of five Dermestid beetles (Coleoptera: Dermestidae) and its implications for phylogeny. *Genomics.* **113**, 927–934 (2021).
20. Huang, J.-P. & Lin, C.-P. Diversification in subtropical mountains: Phylogeography, Pleistocene demographic expansion, and evolution of polyphenic mandibles in Taiwanese stag beetle, *Lucanus formosanus*. *Mol Phylogenet Evol.* **57**, 1149–1161 (2010).
21. Nobrega, M. A. & Pennacchio, L. A. Comparative genomic analysis as a tool for biological discovery. *J Physiol.* **554**, 31–39 (2004).
22. Chen, L. *et al*. Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. *Science.* **364**, eaav6202 (2019).
23. Jaiswal, S. K. *et al*. Genome sequence of peacock reveals the peculiar case of a glittering bird. *Front Genet.* **9**, 392 (2018).
24. Prost, S. *et al*. Comparative analyses identify genomic features potentially involved in the evolution of birds-of-paradise. *GigaScience.* **8**, giz003 (2019).
25. Hu, Y. G., Linz, D. M. & Moczek, A. P. Beetle horns evolved from wing serial homologs. *Science.* **366**, 1004–1007 (2019).
26. Morita, S. *et al*. The draft genome sequence of the Japanese rhinoceros beetle *Trypoxylus dichotomus* septentrionalis towards an understanding of horn formation. *Sci Rep.* **13**, 8735 (2023).
27. Kenzaka, T., Yamada, Y. & Tani, K. Draft genome sequence of an antifungal bacterium isolated from the breeding environment of *Dorcus hopei binodulosus*. *Genome Announc.* **2**, e00424–00414 (2014).
28. Chen, Y., Liu, J., Cao, Y., Zhou, S. & Wan, X. Two new complete mitochondrial genomes of *Dorcus* stag beetles (Coleoptera, Lucanidae). *Genes Genom.* **40**, 873–880 (2018).
29. Huxley, J. S. Relative growth of mandibles in stag-beetles (Lucanidae)*. *Zool J Linn Soc.* **37**, 675–703 (1931).
30. Emlen, D. J. & Nijhout, H. F. The development and evolution of exaggerated morphologies in insects. *Annu Rev Entomol.* **45**, 661–708 (2000).
31. Rowland, J. M. & Emlen, D. J. Two thresholds, three male forms result in facultative male trimorphism in beetles. *Science.* **323**, 773–776 (2009).
32. Liu, B. H. *et al*. *Estimation of genomic characteristicsby analyzing k-mer frequency in de novo genome projects*. Preprint at https://doi.org/10.48550/arXiv.1308.2012 (2013).
33. Guan, D. F. *et al*. Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics.* **36**, 2896–2898 (2020).
34. Hu, J., Fan, J. P., Sun, Z. Y. & Liu, S. L. NextPolish: a fast and efficient genome polishing tool for long-read assembly. *Bioinformatics.* **36**, 2253–2255 (2020).
35. Servant, N. *et al*. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16** (2015).
36. Durand, N. C. *et al*. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).
37. Durand, N. C. *et al*. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst.* **3**, 99–101 (2016).
38. Meyer, J. M. *et al*. Draft genome of the scarab beetle *Oryctes borbonicus* on La Réunion Island. *Genome Biol Evol.* **8**, 2093–2105 (2016).
39. Wang, Q., Liu, L., Zhang, S., Wu, H. & Huang, J. A chromosome-level genome assembly and intestinal transcriptome of *Trypoxylus dichotomus* (Coleoptera: Scarabaeidae) to understand its lignocellulose digestion ability. *GigaScience.* **11**, giac059 (2022).
40. Cunningham, C. B. *et al*. The genome and methylome of a beetle with complex social behavior, *Nicrophorus vespilloides* (Coleoptera: Silphidae). *Genome Biol Evol.* **7**, 3383–3396 (2015).
41. Tribolium Genome Sequencing Consortium The genome of the model beetle and pest *Tribolium castaneum*. *Nature.* **452**, 949–955 (2008).
42. Hoskins, R. A. *et al*. The Release 6 reference sequence of the *Drosophila melanogaster* genome. *Genome research.* **25**, 445–458 (2015).
43. Li, X. Detailed tables for genome assembly and annotations of *Dorcus hopei* (Coleoptera: Lucanidae). *figshare* https://doi.org/10.6084/m9.figshare.24132033.v3 (2023).
44. Xu, Z. & Wang, H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265–W268 (2007).
45. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580 (1999).
46. Tarailo-Graovac, M. & Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinformatics.* **25**, 4.10.11–14.10.14 (2009).
47. Flynn, J. M. *et al*. RepeatModeler2 for automated genomic discovery of transposable element families. *P Natl Acad Sci USA* **117**, 9451–9457 (2020).
48. Bao, W., Kojima, K. K. & Kohany, O. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob DNA.* **6**, 11 (2015).
49. Stanke, M. *et al*. AUGUSTUS: *ab initio* prediction of alternative transcripts. *Nucleic Acids Res.* **34**, W435–W439 (2006).
50. Gertz, E. M., Yu, Y. K., Agarwala, R., Schäffer, A. A. & Altschul, S. F. Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST. *BMC Biol.* **4**, 41 (2006).
51. Birney, E., Clamp, M. & Durbin, R. GeneWise and Genomewise. *Genome Res.* **14**, 988–995 (2004).
52. Kim, D., Landmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* **12**, 357–U121 (2015).
53. Kovaka, S. *et al*. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* **20**, 278 (2019).
54. Haas, B. J. *et al*. Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. *Genome Biol.* **9**, R7 (2008).
55. Camacho, C. *et al*. BLAST plus: architecture and applications. *BMC Bioinform.* **10**, 421 (2009).
56. Mulder, N. & Apweiler, R. InterPro and InterProScan: tools for protein sequence classification and comparison. *Methods Mol Biol.* **396**, 59–70 (2007).
57. Li, X. Genome assembly and annotations of *Dorcus hopei* (Coleoptera: Lucanidae). *figshare* https://doi.org/10.6084/m9.figshare.24123474.v2 (2023).
58. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRP440764 (2023).
59. Li, X. *et al*. Genome assembly AHU_Dhop_1.0. *GenBank* https://identifiers.org/ncbi/insdc.gca:GCA_033060865.1 (2023).
60. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at https://doi.org/10.48550/arXiv.1303.3997 (2013).
61. Danecek, P. *et al*. Twelve years of SAMtools and BCFtools. *GigaScience.* **10**, giab008 (2021).
62. Simao, 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 (2015).
63. Colomba, M. S., Vitturi, R. & Zunino, M. Chromosome analysis and rDNA FISH in the stag beetle *Dorcus parallelipipedus* L. (Coleoptera: Scarabaeoidea: Lucanidae). *Hereditas.* **133**, 249–253 (2000).

Acknowledgements
This work was supported by the National Natural Science Foundation of China (No. 31872276, No. 31572311) and from Yunnan Provincial Science and Technology Department (No. 202105AC160039). We are grateful to Professor Wen Wang (Kunming Institution of Zoology, Chinese Academy of Sciences, and Northwestern Polytechnical University, China) for his support in bioinformatics analysis. We thank Yongjing Chen and Ziqi Li (Anhui University, China) for their assistance in sample collection.

Author contributions
Xueyan Li and Xia Wan conceived and supervised the study. Xiaoyan Bing, Guichun Liu, Zhiwei Dong, and Ruoping Zhao collected samples and sequencing data, and finished relevant experiments. Chuyang Mao and Jinwu He provided analysis guidance. Xiaolu Li and Chuyang Mao performed the analyses. Xiaolu Li wrote the manuscript. Xueyan Li and Xia Wan revised the manuscript. All the authors have read and approved the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to X.W. or Xueyan Li.

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.