

The genome sequence of Philonthus cognatus (Stephens, 1832) (Coleoptera, Staphylinidae), a rove beetle.

Crowley, Liam M.; Telfer, Mark; Geiser, Michael; Mulley, John

DOI: 10.12688/wellcomeopenres.19336.1

Published: 13/04/2023

Early version, also known as pre-print

Cyswllt i'r cyhoeddiad / Link to publication

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA): Crowley, L. M., Telfer, M., Geiser, M., & Mulley, J. (2023). The genome sequence of Philonthus cognatus (Stephens, 1832) (Coleoptera, Staphylinidae), a rove beetle. (pp. 169). (Wellcome open research). Wellcome Open research. https://doi.org/10.12688/wellcomeopenres.19336.1

Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.

You may not further distribute the material or use it for any profit-making activity or commercial gain
 You may freely distribute the URL identifying the publication in the public portal ?

Take down policy If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

The genome sequence of the Rove Beetle, *Philonthus cognatus* (Stephens, 1832)

Authors

Liam M. Crowley, Mark Telfer, Michael Geiser, John F. Mulley, University of Oxford and Wytham Woods Genome Acquisition Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective, Tree of Life Core Informatics collective, Darwin Tree of Life Consortium

Abstract

We present a genome assembly from an individual male *Philonthus cognatus* (the Rove Beetle; Arthropoda; Insecta; Coleoptera; Staphylinidae). The genome sequence is 1,030.6 megabases in span. Most of the assembly is scaffolded into 12 chromosomal pseudomolecules, including the X and Y sex chromosomes. The mitochondrial genome has also been assembled and is 20.7 kilobases in length. Gene annotation of this assembly on Ensembl identified 29,629 protein coding genes.

Keywords

Philonthus cognatus, Rove Beetle, genome sequence, chromosomal, Coleoptera

Liam M.	liam.crowley@biology.ox.ac.uk	Oxford University, Oxford,	Resources,
Crowley		England, UK	Investigation
Mark Telfer	mark.g.telfer@btinternet.com	Independent researcher	Resources,
		Town/City:	Investigation
Michael	m.geiser@nhm.ac.uk	Natural History Museum,	Resources,
Geiser		London, England, UK	Investigation
John F.	j.mulley@bangor.ac.uk	School of Natural Sciences,	Writing –
Mulley		Bangor University, Wales,	original draft
		UK	

Main body

Species taxonomy

Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Staphyliniformia; Staphylinidae; Staphylininae group; Staphylininae; Staphylinini; *Philonthus, Philonthus cognatus* (Stephens, 1832) (NCBI:txid346820).

Background

Philonthus cognatus is a relatively large (8–11 mm), widespread rove beetle (family Staphylinidae), commonly found in damp soils in woodland and grassland habitats across the UK and western Palearctic. It was introduced into North America in the 19th century and is currently present in the USA and Canada (Majka and Klimaszewski, 2008). Adults are typically black in colour, although some individuals may have a bronze or greenish sheen, and the underside of the basal antennal segment is yellow, making members of this species one of the more easily recognised UK rove beetles. Breeding takes place in late spring to early summer and as a result there are two peaks in abundance, with adults most commonly found in the spring and early summer (March to June) and again in autumn (August to October) (NBN Atlas Partnership, 2021), although they are present all year round. *P. cognatus*, like many other rove beetles, are important predators of agricultural pests such as aphids (Dennis and Wratten, 1991; Kollat-Palenga and Basedow, 2000).

The species was named by James Stephens in his multi-volume synopsis of British insects, in the 1832 volume perhaps more famous for referencing some specimens collected by a "C. Darwin Esq." while he was still a student at Cambridge (Stephens, 1832; Barlow, 1958). Stephens described *P. cognatus* as being found "within the metropolitan district, but not common", which seems to contradict their current status as a very common and widespread beetle across the UK and their 'least concern' classification in a recent review of the status of the beetles of Great Britain (Boyce, 2022).

Rove beetles possess defensive glands that produce complex chemical secretions for defence (Parker, 2017; Brückner *et al.*, 2021) or that have antimicrobial activity (Lusebrink, Dettner and Seifert, 2008), and the *P. cognatus* genome assembly will likely prove useful in identifying the biosynthetic pathways involved in the production of these secretions, as well as resolving the polyphyletic status of the species-rich genus *Philonthus* (Chani-Posse *et al.*, 2018).

Genome sequence report

The genome was sequenced from one male *Philonthus cognatus* (Figure 1) collected from Wytham Woods, Oxfordshire, UK (latitude 51.77, longitude –1.34). A total of 29-fold coverage in Pacific Biosciences single-molecule HiFi long reads and 47-fold coverage in 10X Genomics read clouds were generated. Primary assembly contigs were scaffolded with chromosome

conformation Hi-C data. Manual assembly curation corrected four missing or mis-joins and removed two haplotypic duplications, reducing the assembly length by 0.31% and the scaffold number by 2.17%, and increasing the scaffold N50 by 0.67%.

The final assembly has a total length of 1,030.6 Mb in 45 sequence scaffolds with a scaffold N50 of 10.8.9 Mb (Table 1). Most (98.54%) of the assembly sequence was assigned to 12 chromosomal-level scaffolds, representing 10 autosomes, and the X and Y sex chromosomes. Chromosome-scale scaffolds confirmed by the Hi-C data are named in order of size (Figures 2–5; Table 2). While not fully phased, the assembly deposited is of one haplotype. Contigs corresponding to the second haplotype have also been deposited. The mitochondrial genome was also assembled and can be found as a contig within the multifasta file of the genome submission.

The estimated Quality Value (QV) of the final assembly is 55.8 with *k*-mer completeness of 99.99%, and the assembly has a BUSCO v5.3.2 (Manni *et al.*, 2021) completeness of 99.0% (single = 96.8%, duplicated = 2.2%), using the endopterygota_odb10 reference set (n = 2,124).

Genome annotation report

The *Philonthus cognatus* genome assembly GCA_932526585.1 was annotated using the Ensembl rapid annotation pipeline (Table 1; https://rapid.ensembl.org/Philonthus_cognatus_GCA_932526585.1/Info/Index/). The resulting

annotation includes 29,922 transcribed mRNAs from 29,629 protein-coding genes.

Methods

Sample acquisition and nucleic acid extraction

A male *P. cognatus* (icPhiCogn1) was collected from Wytham Woods, Oxfordshire (biological vice-county: Berkshire), UK (latitude 51.77, longitude –1.34) on 8 December 2020. The specimen was taken from woodland habitat by Liam Crowley (University of Oxford) by potting. The specimen was identified by Mark Telfer (independent researcher) and snap-frozen on dry ice.

A second *P. cognatus* specimen (icPhiCogn2) was collected from the Wildlife Garden at the Natural History Museum, London, UK (latitude 51.5, longitude –0.18) on 1 July 2021. The specimen was picked by hand by Michael Geiser (Natural History Museum). The specimen was identified by the collector and preserved at –80°C. The specimen icPhiCogn2 was used for RNA sequencing.

DNA was extracted at the Tree of Life laboratory, Wellcome Sanger Institute (WSI). The icPhiCogn1 sample was weighed and dissected on dry ice with head tissue set aside for Hi-C sequencing. Abdomen tissue was disrupted using a Nippi Powermasher fitted with a BioMasher pestle. High molecular weight (HMW) DNA was extracted using the Qiagen MagAttract HMW

DNA extraction kit. Low molecular weight DNA was removed from a 20 ng aliquot of extracted DNA using the 0.8X AMpure XP purification kit prior to 10X Chromium sequencing; a minimum of 50 ng DNA was submitted for 10X sequencing. HMW DNA was sheared into an average fragment size of 12–20 kb in a Megaruptor 3 system with speed setting 30. Sheared DNA was purified by solid-phase reversible immobilisation using AMPure PB beads with a 1.8X ratio of beads to sample to remove the shorter fragments and concentrate the DNA sample. The concentration of the sheared and purified DNA was assessed using a Nanodrop spectrophotometer and Qubit Fluorometer and Qubit dsDNA High Sensitivity Assay kit. Fragment size distribution was evaluated by running the sample on the FemtoPulse system.

RNA was extracted from thorax tissue of icPhiCogn2 in the Tree of Life Laboratory at the WSI using TRIzol, according to the manufacturer's instructions. RNA was then eluted in 50 μ l RNAse-free water and its concentration assessed using a Nanodrop spectrophotometer and Qubit Fluorometer using the Qubit RNA Broad-Range (BR) Assay kit. Analysis of the integrity of the RNA was done using Agilent RNA 6000 Pico Kit and Eukaryotic Total RNA assay.

Sequencing

Pacific Biosciences HiFi circular consensus and 10X Genomics read cloud DNA sequencing libraries were constructed according to the manufacturers' instructions. Poly(A) RNA-Seq libraries were constructed using the NEB Ultra II RNA Library Prep kit. DNA and RNA sequencing was performed by the Scientific Operations core at the WSI on Pacific Biosciences SEQUEL II (HiFi), Illumina NovaSeq 6000 (RNA-Seq and 10X) instruments. Hi-C data were also generated from head tissue of icPhiCogn1 using the Arima v2 kit and sequenced on the Illumina NovaSeq 6000 instrument.

Genome assembly, curation and evaluation

Assembly was carried out with Hifiasm (Cheng *et al.*, 2021) and haplotypic duplication was identified and removed with purge_dups (Guan *et al.*, 2020). One round of polishing was performed by aligning 10X Genomics read data to the assembly with Long Ranger ALIGN, calling variants with FreeBayes (Garrison and Marth, 2012). The assembly was then scaffolded with Hi-C data (Rao *et al.*, 2014) using YaHS (Zhou, McCarthy and Durbin, 2023). The assembly was checked for contamination and corrected using the gEVAL system (Chow *et al.*, 2016) as described previously (Howe *et al.*, 2021). Manual curation was performed using gEVAL, HiGlass (Kerpedjiev *et al.*, 2018) and Pretext (Harry, 2022). The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2022), which performed annotation using MitoFinder (Allio *et al.*, 2020).

To evaluate the assembly, MerquryFK was used to estimate consensus quality (QV) scores and *k*-mer completeness (Rhie *et al.*, 2020). The genome was analysed and BUSCO scores (Simão *et al.*, 2015; Manni *et al.*, 2021) were generated within the BlobToolKit environment (Challis *et al.*, 2020). Table 3 contains a list of software tool versions and sources.

Genome annotation

The BRAKER2 pipeline (Brůna *et al.*, 2021) was used in the default protein mode to generate annotation for the *Philonthus cognatus* assembly (GCA_932526585.1) in Ensembl Rapid Release.

Ethics and compliance issues

The materials that have contributed to this genome note have been supplied by a Darwin Tree of Life Partner. The submission of materials by a Darwin Tree of Life Partner is subject to the <u>Darwin Tree of Life Project Sampling Code of Practice</u>. By agreeing with and signing up to the Sampling Code of Practice, the Darwin Tree of Life Partner agrees they will meet the legal and ethical requirements and standards set out within this document in respect of all samples acquired for, and supplied to, the Darwin Tree of Life Project. All efforts are undertaken to minimise the suffering of animals used for sequencing. Each transfer of samples is further undertaken according to a Research Collaboration Agreement or Material Transfer Agreement entered into by the Darwin Tree of Life Partner, Genome Research Limited (operating as the Wellcome Sanger Institute), and in some circumstances other Darwin Tree of Life collaborators.

Data availability

European Nucleotide Archive: *Philonthus cognatus* (rove beetle). Accession number <u>PRJEB50787; https://identifiers.org/ena.embl/PRJEB50787</u>. (Wellcome Sanger Institute, 2022)

The genome sequence is released openly for reuse. The *Philonthus cognatus* genome sequencing initiative is part of the Darwin Tree of Life (DToL) project. All raw sequence data and the assembly have been deposited in INSDC databases. Raw data and assembly accession identifiers are reported in Table 1.

Grant information

This work was supported by Wellcome through core funding to the Wellcome Sanger Institute (206194) and the Darwin Tree of Life Discretionary Award (218328).

Author information

Members of the University of Oxford and Wytham Woods Genome Acquisition Lab are listed here: <u>https://doi.org/10.5281/zenodo.4789928</u>.

Members of the Darwin Tree of Life Barcoding collective are listed here: <u>https://doi.org/10.5281/zenodo.4893703.</u>

Members of the Wellcome Sanger Institute Tree of Life programme are listed here: <u>https://doi.org/10.5281/zenodo.4783585</u>.

Members of Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective are listed here: <u>https://doi.org/10.5281/zenodo.4790455</u>.

Members of the Tree of Life Core Informatics collective are listed here: <u>https://doi.org/10.5281/zenodo.5013541</u>.

Members of the Darwin Tree of Life Consortium are listed here: https://doi.org/10.5281/zenodo.4783558.

References

Allio, R. *et al.* (2020) 'MitoFinder: Efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics', *Molecular Ecology Resources*, 20(4), pp. 892–905. Available at: https://doi.org/10.1111/1755-0998.13160.

Barlow, N. (1958) The Autobiography of Charles Darwin 1809-1882. London: Collins.

Boyce, D.C. (2022) A review of the status of the beetles of Great Britain: The Staphylinidae (Rove Beetles) Species Status No. 41. Available at:

http://publications.naturalengland.org.uk/publication/5381819326529536 (Accessed: 14 March 2023).

Brückner, A. *et al.* (2021) 'Evolutionary assembly of cooperating cell types in an animal chemical defense system', *Cell*, 184(25), pp. 6138-6156.e28. Available at: https://doi.org/10.1016/j.cell.2021.11.014.

Brůna, T. *et al.* (2021) 'BRAKER2: Automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database', *NAR Genomics and Bioinformatics*, 3(1), pp. 1–11. Available at: https://doi.org/10.1093/nargab/lqaa108.

Challis, R. *et al.* (2020) 'BlobToolKit – interactive quality assessment of genome assemblies', *G3: Genes, Genomes, Genetics*, 10(4), pp. 1361–1374. Available at: https://doi.org/10.1534/g3.119.400908.

Chani-Posse, M.R. *et al.* (2018) 'Phylogeny of the hyper-diverse rove beetle subtribe Philonthina with implications for classification of the tribe Staphylinini (Coleoptera: Staphylinidae)', *Cladistics*, 34(1), pp. 1–40. Available at: https://doi.org/10.1111/cla.12188.

Cheng, H. *et al.* (2021) 'Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm', *Nature Methods*, 18(2), pp. 170–175. Available at: https://doi.org/10.1038/s41592-020-01056-5.

Chow, W. *et al.* (2016) 'gEVAL — a web-based browser for evaluating genome assemblies', *Bioinformatics*, 32(16), pp. 2508–2510. Available at: https://doi.org/10.1093/bioinformatics/btw159.

Dennis, P. and Wratten, S.D. (1991) 'Field manipulation of populations of individual staphylinid species in cereals and their impact on aphid populations', *Ecological Entomology*, (16), pp. 17–24.

Garrison, E. and Marth, G. (2012) 'Haplotype-based variant detection from short-read sequencing'. Available at: http://arxiv.org/abs/1207.3907.

Guan, D. *et al.* (2020) 'Identifying and removing haplotypic duplication in primary genome assemblies', *Bioinformatics*, 36(9), pp. 2896–2898. Available at: https://doi.org/10.1093/bioinformatics/btaa025.

Harry, E. (2022) 'PretextView (Paired REad TEXTure Viewer): A desktop application for viewing pretext contact maps.' Available at: https://github.com/wtsi-hpag/PretextView (Accessed: 19 October 2022).

Howe, K. *et al.* (2021) 'Significantly improving the quality of genome assemblies through curation', *GigaScience*. Oxford University Press. Available at: https://doi.org/10.1093/gigascience/giaa153.

Kerpedjiev, P. *et al.* (2018) 'HiGlass: Web-based visual exploration and analysis of genome interaction maps', *Genome Biology*, 19(1). Available at: https://doi.org/10.1186/s13059-018-1486-1.

Kollat-Palenga, I. and Basedow, T. (2000) 'Aphid feeding of predatory Staphylinidae on different strata (soil surface and wheat seedlings) in laboratory experiments', *Journal of Plant Diseases and Protection*, (107), pp. 643–648.

Lusebrink, I., Dettner, K. and Seifert, K. (2008) 'Stenusine, an antimicrobial agent in the rove beetle genus *Stenus* (Coleoptera, Staphylinidae)', *Naturwissenschaften*, (95), pp. 751–755.

Majka, C.G. and Klimaszewski, J. (2008) 'Introduced Staphylinidae (Coleoptera) in the maritime provinces of Canada', *The Canadian Entomologist*, (140), pp. 48–72.

Manni, M. *et al.* (2021) 'BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes', *Molecular Biology and Evolution*, 38(10), pp. 4647–4654. Available at: https://doi.org/10.1093/molbev/msab199.

NBN Atlas Partnership (2021) Philonthus cognatus *Stephens, 1832, NBN Atlas*. Available at: https://species.nbnatlas.org/species/NBNSYS0000023442 (Accessed: 20 March 2023).

Parker, J. (2017) 'Staphylinids', Current Biology, (27), pp. R49-R51.

Rao, S.S.P. *et al.* (2014) 'A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping', *Cell*, 159(7), pp. 1665–1680. Available at: https://doi.org/10.1016/j.cell.2014.11.021.

Rhie, A. *et al.* (2020) 'Merqury: Reference-free quality, completeness, and phasing assessment for genome assemblies', *Genome Biology*, 21(1). Available at: https://doi.org/10.1186/s13059-020-02134-9.

Rhie, A. *et al.* (2021) 'Towards complete and error-free genome assemblies of all vertebrate species', *Nature*, 592(7856), pp. 737–746. Available at: https://doi.org/10.1038/s41586-021-03451-0.

Simão, F.A. *et al.* (2015) 'BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs', *Bioinformatics*, 31(19), pp. 3210–3212. Available at: https://doi.org/10.1093/bioinformatics/btv351.

Stephens, J.F. (1832) *Illustrations of British Entomology, Mandibulata, Vol. 5*. London: Baldwin and Cradock.

Uliano-Silva, M. *et al.* (2022) 'MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio High Fidelity reads', *bioRxiv* [Preprint]. Available at: https://doi.org/10.1101/2022.12.23.521667.

Zhou, C., McCarthy, S.A. and Durbin, R. (2023) 'YaHS: yet another Hi-C scaffolding tool', *Bioinformatics*. Edited by C. Alkan, 39(1). Available at: https://doi.org/10.1093/bioinformatics/btac808.

[add dataset citation]

Wellcome Sanger Institute (2022) The genome sequence of the Rove Beetle, *Philonthus cognatus* (Stephens, 1832), European Nucleotide Archive, [dataset], accession number PRJEB50787

Figures



Figure 1. Photograph of the *Philonthus cognatus* (icPhiCogn1) specimen used for genome sequencing.



Figure 2: Genome assembly of *Philonthus cognatus*, **icPhiCogn1.2: metrics.** The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 1,030,568,945 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (181,054,249 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (108,863,099 and 45,099,272 bp), respectively. The pale grey spiral shows the cumulative scaffold count on a log scale with white scale lines showing successive orders of magnitude. The blue and pale-blue area around the outside of the plot shows the distribution of GC, AT and N percentages in the same bins as the inner plot. A summary of complete, fragmented, duplicated and missing BUSCO genes in the endopterygota_odb10 set is shown in the top right. An interactive version of this figure is available at

https://blobtoolkit.genomehubs.org/view/icPhiCogn1.2/dataset/CAKOBO02/snail.



Figure 3: Genome assembly of *Philonthus cognatus*, **icPhiCogn1.2: GC coverage**. BlobToolKit GC-coverage plot. Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/icPhiCogn1.2/dataset/CAKOBO02/blob.



Figure 4: Genome assembly of *Philonthus cognatus*, **icPhiCogn1.2: cumulative sequence**. BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/icPhiCogn1.2/dataset/CAKOBO02/cumulative.



Figure 5: Genome assembly of *Philonthus cognatus*, icPhiCogn1.2: Hi-C contact map. Hi-C contact map of the icPhiCogn1.2 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom.

An interactive version of this figure may be viewed at <u>https://genome-note-higlass.tol.sanger.ac.uk/l/?d=L6nEPxeIQ1KDr06GZSRUKQ</u>.

Tables

Project accession data			
Assembly identifier	icPhiCogn1.2		
Species	Philonthus cognatus		
Specimen	icPhiCogn1		
NCBI taxonomy ID	346820		
BioProject	PRJEB50787		
BioSample ID	SAMEA8603235		
Isolate information	icPhiCogn1.2, abdomen (genome sequencing), head (Hi-C scaffolding)		
	icPhiCogn2, thorax (RNA see	quencing)	
Assembly metrics*		Benchmark	
Consensus quality (QV)	55.8	\geq 50	
k-mer completeness	99.99%	≥95%	
BUSCO**	C:99.0%[S:96.8%,D:2.2%], F:0.2%,M:0.8%,n:2,124	<i>C</i> ≥95%	
Percentage of assembly mapped to chromosomes	98.54%	≥95%	
Sex chromosomes	X and Y	localised homologous pairs	
Organelles	Mitochondrial genome assembled	complete single alleles	
Raw data accessions	·		
PacificBiosciences SEQUEL II	ERR8575390, ERR8575391		
10X Genomics Illumina	ERR8571676–ERR8571679		
Hi-C Illumina	ERR8571680		
PolyA RNA-Seq Illumina	ERR10378003		
Genome assembly			
Assembly accession	GCA_932526585.2		

Table 1: Genome data for *Philonthus cognatus*, icPhiCogn1.2.

Accession of alternate haplotype	GCA_932526485.1	
Span (Mb)	1,030.6	
Number of contigs	131	
Contig N50 length (Mb)	29.5	
Number of scaffolds	45	
Scaffold N50 length (Mb)	108.9	
Longest scaffold (Mb)	181.1	
Genome annotation		
Number of protein-coding genes	29,629	
Number of gene transcripts	29,922	

* Assembly metric benchmarks are adapted from column VGP-2020 of "Table 1: Proposed standards and metrics for defining genome assembly quality" from (Rhie *et al.*, 2021).

** BUSCO scores based on the endopterygota_odb10 BUSCO set using v5.3.2. C = complete [S = single copy, D = duplicated], F = fragmented, M = missing, n = number of orthologues in comparison. A full set of BUSCO scores is available at

https://blobtoolkit.genomehubs.org/view/icPhiCogn1.2/dataset/CAKOBO02/busco.

INSDC accession	Chromosome	Size (Mb)	GC%
OW052241.1	1	181.05	33
OW052242.1	2	148.37	32.7
OW052244.1	3	108.86	32.9
OW052245.1	4	88.74	32.5
OW052246.1	5	67.67	33.2
OW052247.1	6	67.18	32.4
OW052248.1	7	64.79	32.7
OW052249.1	8	55.01	33.6
OW052251.1	9	29.49	35
OW052252.1	10	25.51	36

 Table 2: Chromosomal pseudomolecules in the genome assembly of *Philonthus cognatus*, icPhiCogn1.

OW052243.1	Х	132.81	32.5
OW052250.1	Y	45.1	33.2
OW052253.2	МТ	0.02	28.1
-	unplaced	15.97	35.2

Table 3. Software tools: versions and sources

Software tool	Version	Source
BlobToolKit	4.0.7	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.3.2	https://gitlab.com/ezlab/busco
FreeBayes	1.3.1-17-gaa2ace8	https://github.com/freebayes/freebayes
gEVAL	N/A	https://geval.org.uk/
Hifiasm	0.12	https://github.com/chhylp123/hifiasm
HiGlass	1.11.6	https://github.com/higlass/higlass
Long Ranger ALIGN	2.2.2	https://support.10xgenomics.com/genome- exome/software/pipelines/latest/advanced/other- pipelines
Merqury	MerquryFK	https://github.com/thegenemyers/MERQURY.FK
MitoHiFi	2	https://github.com/marcelauliano/MitoHiFi
PretextView	0.2	https://github.com/wtsi-hpag/PretextView
purge_dups	1.2.3	https://github.com/dfguan/purge_dups
YaHS	1	https://github.com/c-zhou/yahs