nature portfolio

Corresponding author(s): Martin A. Lysak and Christian Parisod

Last updated by author(s): Apr 30, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☒	☐ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Sequencing platforms used to generate the raw data: Illumina HiSeq Xten, Oxford Nanopore Technologies PromethION, and PacBio Sequel II
Data analysis	Genome size measurement: findGSE (v1.94; https://github.com/schneebergerlab/findGSE), Jellyfish (https://github.com/gmarcais/Jellyfish). Genome assembly, scaffolding, and quality assessment: NextDenovo (v2.2; https://github.com/Nextomics/NextDenovo), NextPolish (v1.1.0; https://github.com/Nextomics/NextPolish), hifiasm (v0.19.5; https://github.com/chhylp123/hifiasm), Khaper (https://github.com/lardo/khaper); BUSCO (v3.0.1; https://busco.ezlab.org/), Juicer (v1.5.7; https://github.com/aidenlab/juicer), 3D-DNA (v180922; https://github.com/aidenlab/3d-dna), YaHS (v1.1; https://github.com/c-zhou/yahs), JuiceBox (v1.1.08; https://github.com/aidenlab/Juicebox), GetOrganelle (v1.7.7; https://github.com/Kinggerm/GetOrganelle). Gene prediction and functional annotation: Hisat2 (v2.1.0; https://daehwankimlab.github.io/hisat2/), StringTie (v2.1.4; https://github.com/gpertea/stringtie), GMAP (v2018-07-04; https://github.com/juliangehring/GMAP-GSNAP), TransDecoder (v5.5.0; http://transdecoder.github.io), PASApipeline (v.2.3.3; https://github.com/PASApipeline/PASApipeline), Trinity (v2.11.0; https://github.com/trinityrnaseq/trinityrnaseq), AUGUSTUS (v3.4.0; https://github.com/Gaius-Augustus/Augustus), GeneMark-ET (v4.0; https://genemark.bme.gatech.edu/), GenomeThreader (v1.7.1; https://genomethreader.org/), EVidenceModeler (v1.1.1; https://anaconda.org/bioconda/evidencemodeler), FeatureCounts (https://rnnh.github.io/bioinfo-notebook/docs/featureCounts.html), BLAST (v2.12.0+; https://blast.ncbi.nlm.nih.gov/Blast.cgi), InterProScan (https://github.com/ebi-pf-team/interproscan) Identification of syntenic genes: SynOrths (http://brassicadb.cn:82/download_genome/tools/SynOrths/), NGenomeSyn (https://github.com/hewm2008/NGenomeSyn) Repetitive element annotation: Extensive de novo TE Annotator (EDTA v1.8.3; https://github.com/oushujun/EDTA), LTR_retriever (v2.9.6; https://github.com/oushujun/LTR_retriever), TRASH (https://github.com/vlothec/TRASH), Barrnap (v0.9; https://github.com/tseemann/barrnap), TEsorter (v1.4.6; https://github.com/zhangrengang/TEsorter), soloLTRseeker (https://github.com/estpr/soloLTRseeker), BCT package (https://www.rdocumentation.org/packages/BCT/versions/1.2)

Phylogenetic analysis: MAFFT (v7.427; https://github.com/GSLBiotech/mafft), TrimAL (v1.4; https://vicfero.github.io/trimal/), IQ-TREE (v1.6.11; http://www.iqtree.org/), ASTRAL-Pro (v1.1.3; https://github.com/chaoszhang/A-pro), Toytree (v.3.0.5; https://github.com/eaton-lab/toytree)
Subgenome phasing and validation: SubPhaser (v1.2; https://github.com/zhangrengang/SubPhaser).
Calibration of the WGD timing: KaKs_Calculator (v2.0; https://github.com/lizzhao/Kaks_Calculator), r8s (v1.81; https://sourceforge.net/projects/r8s/), PATHd8 (v1.0; https://www2.math.su.se/PATHd8/), RelTime method in MEGA X (https://www.megasoftware.net/), OrthoFinder (v2.5.4, https://github.com/davidemms/OrthoFinder),
Reconstruction of the ancestral karyotype: WGDI (https://github.com/SunPengChuan/wgdi)
Gene functional enrichment: clusterProfiler package (v3.14.3; https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html)
3D genome analysis: BWA-MEM (v0.7.17; https://github.com/lh3/bwa), HiCExplorer (v3.7.2; https://github.com/deeptools/HiCExplorer)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The raw genome sequencing data generated in this study have been deposited in China National GeneBank (CNGBdb; https://db.cngb.org/cnsa/) under the accession CNP0007292. The genome assembly and annotation data have been deposited in Figshare (https://doi.org/10.6084/m9.figshare.28451828.v1). All other data needed to evaluate the conclusions in the manuscript are present in the paper and/or the Supplementary Information.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were required to establish sample size for this study. To cover different genome sizes and chromosome numbers, 8 representative Biscutella species were chosen for genome sequencing.
Data exclusions	No samples were excluded. Raw sequencing data were quality filtered as described in the manuscript.
Replication	All experiments were performed at independent experimental days and showed similar tendencies. The genome assemblies and annotations were validated using multiple methods. Other replication analyses that support the conclusions are indicated in the manuscript, e.g., phylogenomic analyses by multiple methods, subgenome phasing by multiple methods.
Randomization	Samples were randomly allocated into experimental groups.
Blinding	Blinding not applicable for genome sequencing and assembly.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☐	☒ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks: The following Biscutella accessions were used: B. laevigata subsp. varia (V12-4; Germany, Beuron), B. laevigata subsp. austriaca (Jord.) Mach.-Laur. (A2Schnee 3B; Austria, Schneealpe Altenberg), B. prealpina Raffaelli & Baldoin (RCBO_NC17; Italy, Recoaro Terme), B. frutescens Coss. (PI 650129, USDA collection; Spain), B. auriculata L. (PI 650127, USDA collection; Spain, Ames), B. didyma L. (LBN-08579, LBN seed bank; Lebanon), B. baetica Boiss. & Reut. (Gaucín; Spain, Gaucín), and B. lyrata L. (Cádiz; Spain, Cádiz).

Novel plant genotypes: No novel genotypes were produced.

Authentication: No authentication procedures were used.