nature research

Corresponding author(s): Jianquan Liu

Last updated by author(s): May 19, 2021

Reporting Summary

Nature Research 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 Research 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	No software had been used for data collection
Data analysis	Software used are listed as follows:ncbi-BLAST+ (v2.2.28), FALCON (Branch 3.1), Pilon (v1.22), Quiver (smrtanalysis v3.0.2) , LACHESIS (v1), BUSCO (v3), NGS QC Toolkit (v2.3.3), Trinity (v2.4.0), RepeatMasker (v4.0.7), RepeatModeler (v1.0.10), LTR Finder (v1.02), LTRharvest (genometools v1.5.9), LTR_retriever (v2.6), AUGUSTUS (v3.2.2), PASA pipeline (v2.1.0), HISAT2 (v2.1.0), InterProScan (v5.25-64.0), Blast2GO (v4.1), OrthoMCL (v2.0.9), Cafe (v3.1), Cufflinks (v2.2.1), MAFFT (version 7), RAxML (v7.2.3), PAML (v4.9e), HMMER (v 3.1b2), MCL (version 14-137), MCScanX (v1), LAST (v963), Cytoscape (v3.6.0), AutoDock (v4.2.6), R (v3.4.3), Jellyfish (v2.2.9), PyMOL (v2.2.0).

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 Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

The PacBio long reads and Illumina short reads have been deposited in the NCBI Sequence Read Archive (SRA) database under BioProject PRJNA639006 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA639006). The final chromosome-scale genome assembly and the GFF3 file have been deposited in Figshare (https://doi.org/10.6084/m9.figshare.12570599 and https://doi.org/10.6084/m9.figshare.12570614). The SwissProt and TrEMBL database used in this study are available from https://www.uniprot.org. KEGG Pathway database can be available from https://www.kegg.jp.

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	We sequenced a single Camptotheca acuminata plant, and no statistical methods were used to predetermine sample sizes.
Data exclusions	No data were excluded from the analyses.
Replication	The genome sequence was taken and sequenced with more than 100 fold coverage. The LAMT enzyme activity assay experiments were replicated for 3 time, all attempts at replication were successful. The SDS-PAGE experiments were repeated independently for 3 times with similar results.
Randomization	No random sampling is required for genome sequencing, because the genome differences are very small within the wild population, thus any wild plant is allowed for genome sequencing.
Blinding	Blinding is not applicable in our study because it does not involve subjects which receive different treatments.

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
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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