nature portfolio

Corresponding author(s): Yongfeng Zhou

Last updated by author(s): Jan 28, 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	No software had been used for data collection
Data analysis	HiFiasm v0.16.1-r375; GenomeScope v2.0; Juicer v1.5; 3D-DNA ; Juicebox v1.11.08; NextDenovo ; Minimap2 v2.24; IGV v2.12.3; BUSCO ; Hisat2 v2.10.2; StringTie v1.3.0; AUGUSTUS v3.4.0; RepeatModeler open-2.0.3; RepeatMasker open-4.1.2; TIDK v0.2.0; TRF2GFF ; Plotsr ; MUMmer v.4.0; VCFtools v0.1.16; Fastp v0.21; BWA v 0.7.15; SAMtools v 1.4; GATK v4.1.8; PLINK v1.90b6.21; FastTree v2.1.10; Admixture v1.3.0; SweeD v3.2.1; Beagle v5.4; superSFS ; SIFT 4G ; Swiss-Prot ; ANGSD ; Delly v1.0.3; BCFtools v1.13; Bedtools ; Orthofinder v2.5.4; IQ-TREE v2.1.4; Last ; GERP; clusterProfiler v4.0; diamond; popgenWindows.py

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

All sequence data have been deposited in the NCBI Sequence Read Archive under project number PRJNA951461 and the National Genomics Data Center (NGDC) Genome Sequence Archive (GSA) (https://ngdc.cncb.ac.cn/gsa/),with BioProject number PRJCA016741. The assembly and annotation have been deposited in zenodo: https://zenodo.org/record/8080252. VCF files for SNPs and SVs are avaliable from the authors on request.

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	*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used.* *Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected.* *Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*
Reporting on race, ethnicity, or other socially relevant groupings	*Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status).* *Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.)* *Please provide details about how you controlled for confounding variables in your analyses.*
Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

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

Ecological, evolutionary & environmental sciences study design

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

Study description	In this study, we study the population genomics of diverse V. vinifera germplasm sets that span various mating histories, ranging from selfing to obligately outcrossing. Our study explores the evolutionary genomic mechanisms underlying the transitions of mating systems in forming the Pinot Noir lineages, which shed insights on genomic breeding of grapevine.
Research sample	Our population samples include wild grapes (V. vinifera subsp. sylvestris), domesticated grapes Pinot Noir, Schiava Grossa, Gouais Blanc, Chardonnay, Gamay Noir, Helfensteiner, and PN40024.
Sampling strategy	Fresh and healthy leaf tissue from plants of Vitis vinifera L. 'Pinot Noir' was collected from Shenzhen, Guangdong Province, China, and immediately frozen in liquid nitrogen. These materials were packaged for PacBio HiFi, ONT ultra-long, and Hi-C sequencing, respectively, and subsequently submitted to the company for genomic DNA extraction and library preparation.

Data collection	50 grape samples were collected for analysis, of which 9 Pinot Noir were performed in-house and the rest were all downloaded from the National Center for Biotechnology Information (NCBI).
Timing and spatial scale	The Pinot Noir sample for assembly was sampled in 2021，the rest of the grapes are obtained from public sources。
Data exclusions	No data were excluded.
Reproducibility	All attempts to repeat the analyses were successful.
Randomization	Not applicable.
Blinding	The generation of sequencing data was blind.

Did the study involve field work? ☐ Yes ☒ No

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

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes
☒	☐ Public health
☒	☐ National security
☒	☐ Crops and/or livestock
☒	☐ Ecosystems
☒	☐ Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☒	☐	Demonstrate how to render a vaccine ineffective
☒	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☒	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☒	☐	Increase transmissibility of a pathogen
☒	☐	Alter the host range of a pathogen
☒	☐	Enable evasion of diagnostic/detection modalities
☒	☐	Enable the weaponization of a biological agent or toxin
☒	☐	Any other potentially harmful combination of experiments and agents

Plants

Seed stocks	The Cultivar Pinot Noir clone (AGIS_01) for genome assemblies was collected from the grapevine germplasm collection in the Agricultural Genomics Institute at Shenzhen, Shenzhen, Guangdong Province, China. The nine clonal Pinot Noir for resequencing were from Foundation Plant Services, University of California, Davis, USA.
Novel plant genotypes	No novel plant genotypes
Authentication	The plant stock has authentication by the Agricultural Genomics Institute at Shenzhen, Shenzhen, Guangdong Province, China