

## 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                                                                                                                                                                                                                                                                                      |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement                                                                                                                                    |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly                                                                                                                                    |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>                                                               |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of all covariates tested                                                                                                                                                                                                                     |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons                                                                                                                                        |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> 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) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings                                                                                                                                                                      |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes                                                                                                                                                |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> 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 was used for data collection                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Data analysis   | guppy v5.0.16, minimap2 v2.24, bwa-mem2 v2.2.1, longshot v0.4.5, gatk v4.5.0.0, gatk v4.2.3.0, plink v1.9, whatshap phase v1.4, filtlong v0.2.1, samtools v1.15.1, samtools v1.16.1, bcftools v1.18, bcftools v1.16, Necat v0.0.1_update20200803, Flye v2.9, SMARTdenovo, medaka v1.8.0, HapoG v1.3.3, ragout v2.3, Merqury v1.3, miniBusco v0.2.2, LRSDAY pipeline v1.7.0, Telofinder, MUM&Co v3.8, MUMmer4, Jasmine v1.1.5, blast v2.12.0, Python package NetworkX v3.2, SnpEff v5.1, InterProScan v5.65-97.0, STAR v2.7.9a, R package ape v5.8, R package SNPRelate v1.26.0, R package chisq.posthoc.test v0.1.2, R package iNEXT v3.0.0, R package fgsea v1.27.0, R package rrvgo v1.10.0, LDAK v5.2, Python package FaST-LMM v0.4.6, Minigraph-Cactus pipeline v2.6.4, vg toolkit v1.54.0, Python v3.11.5, Python v3.8.10, R v4.3.1 |

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](#)

Sequencing reads associated with this work are available at the European Nucleotide Archive under the accession no. PRJEB77686 and PRJEB81147. Genomes and annotations, gene-based pangenome, graph-based pangenomes, SV matrix and phenotypes generated are available in a Zenodo repository (<https://doi.org/10.5281/zenodo.15698884>).

## 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                                        | Not applicable |
| Reporting on race, ethnicity, or other socially relevant groupings | Not applicable |
| Population characteristics                                         | Not applicable |
| Recruitment                                                        | Not applicable |
| Ethics oversight                                                   | Not applicable |

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](https://www.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     | Sample size is n=1,086, which mostly overlap samples for which transcriptomic, proteomic and growth phenotypes were previously generated (Peter et al., 2018, Nature ; Caudal et al., 2024, Nature genetics ; Teyssonniere et al., 2024, PNAS ; Muenzner et al., 2024, Nature). |
| Data exclusions | Samples with low sequencing depth were discarded from the analyses.                                                                                                                                                                                                             |
| Replication     | No sequencing replicates were performed, because correspondence with previously generated sequencing data (Peter et al., 2018, Nature) was confirmed.                                                                                                                           |
| Randomization   | Isolates with similar genetic backgrounds were randomly spread across multiple sequencing batches.                                                                                                                                                                              |
| Blinding        | Blinding was not relevant to this study because the research design did not involve a comparison that could introduce bias.                                                                                                                                                     |

## 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                                 | Involvement in the study                               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                 |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Plants                        |

## Methods

|                                     |                                                 |
|-------------------------------------|-------------------------------------------------|
| n/a                                 | Involvement in the study                        |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Plants

Seed stocks

Not applicable

Novel plant genotypes

Not applicable

Authentication

Not applicable