nature portfolio

Corresponding author(s): Neuveglise Cécile

Last updated by author(s): Jun 1, 2026

# 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<br>*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<br>*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 | The Method section of the manuscript lists all the codes and software used in this study, along with their versions. All codes have been released and deposited in Zenodo, a repository that assigns Digital Object Identifiers (DOI: https://doi.org/10.5281/zenodo.17613032). The code is also available on an institutional forge, where it can be cloned (https://forge.inrae.fr/frederic.bigey/diutina-catenulata-code.git). |
| Data analysis | The same codes and software described above were used for data analysis. |

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 data supporting the findings of this study are available within the manuscript and the supplementary tables. The raw sequencing data and annotated genomes were deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under the umbrella project accession number PRJEB94530. The corresponding accession numbers (Project, BioSample, and Run) for each strain are available in Supplementary Table 1.

## 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 | As this was a diversity survey, no sample size calculation was performed. |
| Data exclusions | No data were excluded from the analyses. |
| Replication | No specific measures were taken to ensure reproducibility. The 61 yeast strains were only sequenced once. The kayotypes were also only done once, except for the reference strains. The ploidy of the 61 strains was estimated once using a flow cytometer. |
| Randomization | Randomisation was not applicable to this study. |
| Blinding | Blinding was not relevant to this study. |

# 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 |

## Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

| | |
|---|---|
| Laboratory animals | n/a |
| Wild animals | n/a |

| | |
|---|---|
| Reporting on sex | n/a |
| Field-collected samples | The yeasts were collected from various origins and maintained as described in the Methods section of the manuscript. |
| Ethics oversight | No ethical approval or guidance was necessary for working with yeast cells. |

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

## Plants

| | |
|---|---|
| Seed stocks | n/a |
| Novel plant genotypes | n/a |
| Authentication | n/a |

## Flow Cytometry

### Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

| | |
|---|---|
| Sample preparation | Yeast cells were prepared for flow cytometry analysis as previously described (Saubin et al., 2019). Briefly, cells were grown overnight in YPD medium at 28 °C under 180 rpm agitation, and then diluted to OD600 0.1 in 10 mL YPD and placed in the same conditions during 5 h. After OD measuring, about 107 cells were centrifuged 1 min at 10,000 rpm with 1 mL water. The cell pellets were then suspended in 1 mL water, added drop by drop in 8 mL of ethanol 75% with permanent vortexing and finally stored at 5 °C. After one night at 5°C, the cells were centrifuged 5 min at 3,000 rpm, suspended in 1 mL of PBS buffer and centrifuged once again 1 min at 13,000 rpm. Cell pellets were then suspended in 500 mL of RNase A (2 mg/mL in 10 mM Tris-Cl and 15 mM NaCl) and incubated 1 h at 37 °C. Finally, cells were treated 1 h at 50 °C with 200 mL of 1 mg/mL of proteinase K diluted in PBS buffer, centrifuged and suspended in 500 mL of PBS buffer before being sonicated 15 s in a Branson Sonifier 250 sonicator at 10% of the maximum power. About 106 cells were labeled with SYTOXR green (Invitrogen) in 250 mL at a final concentration of 1 mM. DNA labeled with SYTOX R green was quantified on a C6 Accuri (AnnArbor, MI, United States) flow cytometer with an excitation wavelength of 488 nm and an emission wavelength of 530 nm. Acquisition was performed on 30,000 events observed with gating on forward scatter/side scatter signals. The flow rate was set to about 2,000 events per second (medium flow, 35 mL/min;core, 16 mm). |
| Instrument | Accuri C6 (Ann Arbor, MI, United States) flow cytometer |
| Software | BD Accuri C6 software version 1.0.264.21 |
| Cell population abundance | Acquisition was performed on 30,000 cells with a flow rate set to about 2,000 events per second. |
| Gating strategy | Cells were analyzed by gating on forward scatter (FSC) and side scatter (SSC) signals, and an initial gate was defined to identify yeast cells. Gates were then established on forward scatter versus fluorescence intensity plots according to the ploidy levels of the reference strains. Finally, the ploidy of the studied strains was determined directly in the software by comparison with the reference strains (Source data file 2). |

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.