

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
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☒ | ☐ | 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
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ | 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 paired reads were first merged with fastp, before being converted from fastq to fasta format with seqtk (). A blast database was generated from all these files and queried with the blastn program from a fasta file containing 538 non-ambiguous barcodes related to strains with the following parameters: -word_size 13; -strand plus; -dust no. Number of occurrence of each barcode was extracted from the blast result files.

Data analysis

All analyses were performed using R (4.5.1). The following packages were used: pheatmap (1.0.13); dendextend (1.19.1); stats (4.5.1); fgsea (1.34.2). GWAS was performed using FaST-LMM (0.6.4). Codes available at <https://github.com/HaploTeam>.

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 barcoded strain collection is available upon request. Sequence data are available from National Centre for Biotechnology Information Sequence Read Archive

under accession number: PRJEB96750.

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	NA
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	NA
Recruitment	NA
Ethics oversight	NA

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	We barcoded 538 <i>Saccharomyces cerevisiae</i> isolates derived from 1011 natural isolate collection (Peter et al. 2018, Nature)
Data exclusions	Conservative quality thresholds (edit distance >4 and GC content between 30% and 70%) yielded a high-confidence barcoded collection of 520 isolates. Finally, a master pool of 495 barcoded isolates was created after excluding strains with genome-encoded killer viruses or flocculation phenotypes to prevent confounding interactions. Using the barcoded collection, we screened a library of 633 natural compounds from the Riken natural product depository. Following initial screening, 178 compounds were excluded. These included 120 compounds (~19%) that showed complete growth inhibition after 24 hours and 58 compounds that did not pass our data quality control criteria.
Replication	Biological replicates were done for the benchmarking experiment. All replicate data was successfully obtained.
Randomization	Equal volumes of independent saturated cultures of the 495 barcoded isolates were pooled after 48 hours. This pool was inoculated into 200 uL YPD media supplemented with compounds and grown for 24 hours in 96-well plates.
Blinding	All analyses were performed without preconception of the ecological and geographical origins, domestication status or genetic similarity between the samples.

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	NA
Novel plant genotypes	NA
Authentication	NA