nature portfolio

Corresponding author(s): Olivier Tenaillon

Last updated by author(s): Jan 20, 2022

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	All data analysis pipeline will be made available in a Dryad repository when the article will be in its final stages. For the analysis of the raw data, an html Mark down file has been made that describes all the command line used to go from the fastq files to a table of barcodes counts with an assigned sequence. Overall Mothur (https://www.mothur.org/) R and python were used.
Data analysis	For the inference of the fitness from the counts of the barcodes, we have developed our own inference procedure. The theory is presented in the file "supplementary_material_inference.pdf". The implementary of the inference procedure can be find in inference_procedure.py and inference_formula.py A jupyter notebook, inference.ipynb, builds on the two previous python files, and shows our inference procedure. Starting from the barcode counts, running the different cells of the notebook, we find the fitness. These three files are located in the folder Data/Inference.

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 the sequencing data will be made accessible through the European Nucleotide Archive.

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 we expected to cover about 20000 genotypes, we were satisfied with mutant library of 150000 mutants. From this pool of mutants about 2 million barcoded mutants were generated, such that roughly each mutant will be associated to 10 different barcodes. Overall this strategy allowed us to have robust enough data for about 15000 genotypes.
Data exclusions	All procedures leading to some filtering of the data have been described in the methods. Basically, genotypes with too low number of barcodes, or with too low total read counts were excluded. Because the assignation of genotype to barcode can be affected by recombination, for each genotypes, barcodes with divergent pattern were excluded as explained in the material and methods.
Replication	A duplicate has been made, in which the library of mutants was exposed independently to the same antibiotic selective pressure. Because the control of antibiotic concentration is difficult and our resutls are sensitive to this concentration, both experiments were made on the same day.
Randomization	na
Blinding	na

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