nature portfolio

Corresponding author(s): Pouria Dasmeh

Last updated by author(s): May 17, 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	We used the R library flowCore (v.2.8.0), and the software FlowJo V10.4.2 (LLC) to analyze flow cytometry data.
Data analysis	We performed all statistical analyses with R (v.4.0) and employed the Bioconductor package (v.3.15).

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 datasets generated, raw and processed files, as well as scripts are available at: https://github.com/dasmeh/Trait_correlation

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 measured the correlation between the two traits of yellow fluorescent protein (YFP), namely the ability to emit yellow and green color. We performed experiments in populations of E.coli cells that expressed variants of YFP. Our experiments consisted of fluorescent measurements in live cells, as well as in vitro measurements of protein stability, foldability, and solubility as described in Methods.
Research sample	The samples included 1) E. coli cells that express yellow fluorescent proteins, and 2) cell lysates for biochemical measurements. All measurements were done with three different biological replicates. In vitro experiments were done with three technical replicates.
Sampling strategy	We randomly sampled 90 single clones per subpopulations studied in the work.
Data collection	We used Aria III cell sorter (BD Biosciences) to sort cells. We used a TECAN microplate reader (TECAN Spark) to measure yellow and green fluorescent intensities.
Timing and spatial scale	No spatial or temporal resolution was needed in these experiments.
Data exclusions	For experimental tractability we selected 10 YFP variants that included one of the neo-functionalizing mutations G55S, and Y204C and their double mutants. We excluded the rest of mutants for this selected study but then used the full polymorphic subpopulations of YFP to measure protein stability, foldability, and solubility.
Reproducibility	All protocols are explained in the methods. We are also happy to provide aliquots/samples of our subpopulations to interested academic parties.
Randomization	No randomization was needed for this study.
Blinding	We repeated our experiments in replicates but no blinding procedure was employed in this study.

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
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

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	We sampled 40 ul of cell cultures and suspended it in 2ml cold PBS buffer as explained in Methods.
Instrument	Aria III cell sorter (BD biosciences)
Software	R library flowCore (v.2.8.0), and the software FlowJo V10.4.2 (LLC)
Cell population abundance	We collected 10^5 cells in ~ 1ml LB medium for each of our subpopulations.
Gating strategy	We measured the yellow fluorescence in the FITC channel (lambda_ex=488 nm, and lambda_em=530 nm), and green fluorescence in the AmCyan channel (lambda_ex=405 nm, and lambda_em=525 nm). We will add the figure on gating strategy to the supplementary information upon revision.

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