

Corresponding author(s): Guan Wang, Hongzhong Lu

Last updated by author(s): Apr 11, 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
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 The whole dataset is available in Github repository (<https://github.com/hongzhonglu/EnzymeTuning>).

Data analysis All code for the analysis are included in the repository (<https://github.com/hongzhonglu/EnzymeTuning>).
The GAN framework uses Python based code, the version for the python packages:

Python v3.9.19
Numpy v1.26.4
Pandas v2.3.1
cobra v0.29.1
seaborn v0.13.2
Matplotlib v3.5.2
plotly v6.1.1
scikit-learn v1.6.1
scipy v1.13.1
gurobipy v11.0.3
cplex v22.1.1.0
h5py v3.12.1
tensorflow v2.16.2
keras v3.5.0

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 proteome raw data have been uploaded to the iProX database [<https://www.iprox.cn/page/home.html>]. Project ID is IPX0015844000, and ProteomeXchange ID is PXD074934. The data in this study are provided in the source data file and in the GitHub repository [<https://github.com/hongzhonglu/EnzymeTuning>].

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	The sample size in this study was determined to be 237 kcat parameters. We initially performed a global sensitivity analysis using the SALib to quantify the influence of all 4,354 kcat parameters in the ecGEM on growth phenotype predictions. We identified kcat parameters that were insensitive to growth phenotype prediction and were intersected with the set of enzymes associated with extreme protein abundance predictions.
Data exclusions	No data were excluded.
Replication	All experiments in this study were performed independently with biological triplicates, and the data are presented as the mean $\pm$ standard deviation.
Randomization	Not applicable.
Blinding	Not applicable.

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
- ☒

☐ Antibodies
- ☒

☐ Eukaryotic cell lines
- ☒

☐ Palaeontology and archaeology
- ☒

☐ Animals and other organisms
- ☒

☐ Clinical data
- ☒

☐ Dual use research of concern
- ☒

☐ Plants

Methods

- n/a

Involvement in the study
- ☒

☐ ChIP-seq
- ☒

☐ Flow cytometry
- ☒

☐ MRI-based neuroimaging

Plants

Seed stocks

Not applicable.

Novel plant genotypes

Not applicable.

Authentication

Not applicable.