nature portfolio

Corresponding author(s): YANG MENG

Last updated by author(s): Jun 10, 2023

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 Python(3.8) to process and collect the data in this study.
Data analysis	We have made the code available on GitHub. The repository can be found at https://github.com/melobio/EvoPlay The trained weights of TAPE Oracle are obtained from https://github.com/HeliXonProtein/proximal-exploration. The trained weights of ESM (1b,1v) models for preselecting action sites are obtained from https://github.com/facebookresearch/esm. The implementation of benchmarking methods(Evolutionary-BO, Cbas, DyNA-PPO, Adalead, SAC) are based on FLEXS library https://github.com/samsinai/FLEXS. We use AlphaFold2(2.0) as surrogate predictor for the design of peptide-receptor interface structure. The modification of EvoBind with recycle 0 and 3 are based on https://github.com/patrickbryant1/EvoBind. The implementation of MCMC in peptide design are based on the repository https://github.com/gjoni/trDesign. The benchmarking method of CLADE for generating 384 training sequences for ftMLDE supervised learning is as the implementation of https://github.com/WeilabMSU/CLADE. For generating the 96 sequences of 2nd stage of EvoPlay-assisted MLDE, we use the supervised learning of ensemble models from https://github.com/fhalab/MLDE. We utilize DeepMSA v1 pipeline service (https://zhanglab.ccmb.med.umich.edu/DeepMSA/) to obtain a 30-sequence multiple sequence alignment (MSA) and use Weblogo 2.8.2(https://weblogo.berkeley.edu/logo.cgi) to generate sequence logo for Gaussia luciferase (GLuc). The implementation of UniRep in ablation experiments of surrogate models are based on the repository https://github.com/churchlab/UniRep .

The implementation of Hybrid in ablation experiments of surrogate models are based on the repository https://github.com/Protein-Engineering-Framework/Hybrid_Model.
The implementation of MuFacNet in ablation experiments of surrogate models are based on the repository https://github.com/HeliXonProtein/proximal-exploration.

We adopt the following software for Molecular Dynamic Simulation and analysis: Gromacs (2020.1),Sobtop (dev3.1),ORCA (5.0),PyMOL(2.1), Autodock (4.2).

The main algorithms related to molecular simulation include the following:
Lamarckian algorithm,
Particle Mesh Ewald (PME) algorithm
Steepest descent algorithm
Velocity rescaling algorithm
Parrinello-Rahman barostat algorithm
Linear constraint solver (LINCS) algorithm
UniDesign

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

Datasets used in all benchmark studies were published previously.

The PAB1 (UniProt ID: P04147) and GFP (UniProt ID: P42212) datasets was obtained from https://doi.org/10.6084/m9.figshare.23498195.

Peptide and receptor sequences of the wild type for 1ssc,2cnz,3r7g and 6seo (PDB id) can be referred to FigShare (doi: https://doi.org/10.6084/m9.figshare.23375666.v1)

GB1 (UniProt ID: P19909) and PhoQ (UniProt ID: P23837) datasets were obtained from https://doi.org/10.6084/m9.figshare.21767369.v3.

The MSA and logo used in GLuc design is presented in supplementary data 1 and Extended Data Fig. 1.

Our in-house GLuc library can be referred to Supplementary Data 2.

Experimentally validated variants proposed by EvoPlay are presented in Supplementary Data 3

Source Data are provided by the link: https://doi.org/10.6084/m9.figshare.23437295.v1

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	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	The sample sizes were sufficient for analysis, and no statistical methods were used to predetermine sample size.
Data exclusions	1) For GLuc experiment, samples with mutating sites within the loosely packed moiety (sites 150-168) of GLuc were excluded during in-house pool construction. Samples that were not successfully constructed or purified were excluded. 2) For bio-layer interferometry validation test, no data were excluded.
Replication	1)For the GLuc experiments, results were reproduced twice independently using different batches of purified proteins on separated days, and for each independent experiment, bioluminescence measurements were repeated three times successfully with similar results. 2)Biolayer interferometry experiment were independently repeated at least twice, and data were collected in triplicate over several days. All attempts were successful. The protein–peptide interaction kinetics were assessed at various concentrations. All data were reproducible.
Randomization	Samples were randomly separated into experimental groups.
Blinding	All experiments were performed by researchers unaware of the identities of the samples(researchers unaware of the predicted fitness score predicted by models and even which model a sample was generated).

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

Methods

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