

Corresponding author(s): Min LiLast updated by author(s): Feb 11, 2025

## 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

- |                                     |                                     |                                                                                                                                                                                                                                                            |
|-------------------------------------|-------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement                                                                                                                                    |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly                                                                                                                                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>                                                               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | A description of all covariates tested                                                                                                                                                                                                                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons                                                                                                                                        |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | 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) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings                                                                                                                                                           |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes                                                                                                                                     |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | 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 | <input type="text" value="No software was used for data collection."/>                                                                                  |
| Data analysis   | <input type="text" value="Data are analyzed using numpy (1.21.2), pandas (1.3.5) and Matplotlib (3.5.1). Structures are visualized by Pymol (2.5.2)."/> |

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 of data information can be obtained from the "Data Availability" Section in the manuscript. The data of this work are available at <https://github.com/CSUBioGroup/DTIAM>.

## 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                                        | n/a |
| Reporting on race, ethnicity, or other socially relevant groupings | 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](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 evaluate our method on the six benchmarks and the sample size of the six benchmarks (Yamanishi_08, Hetionet, Davis, KIBA, Activation and Inhibition) is 56397, 549362, 30056, 118254, 21043 and 231605, respectively. For binary prediction of DTI and MoA, the sample size is determined by the number of positive samples in the dataset as well as the ratio of positive and negative sample. For regression prediction of DTA, the sample size is determined by the number of known affinity labels in the dataset. Basic statistics of six benchmark datasets are available in the Supplementary Information. |
| Data exclusions | No data were exclude.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Replication     | All source code and benchmark datasets are available, the results are reproducible.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| Randomization   | For the test experiments, data allocation to training and test sets was performed randomly using a K-fold cross-validation approach. Each dataset is randomly split into K folds, and the model is trained K times, where each time a different fold is held out as the test set, and the remaining K-1 folds are used as the training set.                                                                                                                                                                                                                                                                           |
| Blinding        | For the test experiments, data allocation to training and test sets was performed randomly using a K-fold cross-validation approach. The investigators were blinded to group allocation during experiments and outcome analysis.                                                                                                                                                                                                                                                                                                                                                                                      |

## 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                                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                       |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms      |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Plants                           |

### Methods

|                                     |                                                 |
|-------------------------------------|-------------------------------------------------|
| n/a                                 | Involved in the study                           |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

|                                                                      |                                                                                                                       |
|----------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|
| Cell line source(s)                                                  | HEK293T cells and LA795 cells were purchased from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China). |
| Authentication                                                       | As these are commercially available cell lines, no further cell line authentication was performed.                    |
| Mycoplasma contamination                                             | The cell lines were tested and are free from mycoplasma contamination.                                                |
| Commonly misidentified lines<br>(See <a href="#">ICLAC</a> register) | No commonly misidentified cell lines were used in the study.                                                          |

## Plants

|                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|-----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Seed stocks           | <i>Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.</i>                                                                                                                                                                                                                                                                                          |
| Novel plant genotypes | <i>Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.</i> |
| Authentication        | <i>Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.</i>                                                                                                                                                                                                                                       |