nature portfolio

Corresponding author(s): Daniel Merk

Last updated by author(s): Jul 29, 2024

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	Compound-bioactivity data were collected from public resources (ChEMBL, PubChem, IUPHAR/BPS, BindingDB, Probes&Drugs) and combined using KNIME (v4.5) with rdkit software (version 2022.09.1).
Data analysis	Data were processed and analyzed using KNIME (version 4.5), Python (version 3.7 (CLMs) and 3.9 (data analysis and visualization)), Python module scikit-learn (version 1.2.2), tensorflow (version 2.14.0), rdkit (version 2022.09), matplotlib (version 3.8.4), keras (version 2.14.0), pandas (version 2.2.2), numpy (version 1.26.4), CATS descriptors (https://github.com/iwatobipen/CATS2D), DataWarrior (version 5.5), MOE (version 2022.02). Code used in this study is available at Zenodo (doi: 10.5281/zenodo.12795470).

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 data generated in this study and supporting the results of this are provided in the Supplementary Information and Source Data file. Code used in this study is available at Zenodo (doi: 10.5281/zenodo.12795470)

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

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Only in vitro assays have been performed. Sample sizes for in vitro assays were determined to achieve at least R^2>0.95 for dose-response curves and were always at least n=3.
Data exclusions	No data have been excluded
Replication	Only in vitro assays have been performed. All in vitro assays have been successfully replicated in at least three biologically independent repeats.
Randomization	Not relevant in this study. Only in vitro assays were conducted in which samples are derived from the same clone. Different repeats were conducted on different dates and from different harvests/transfections/seeds to account for covariates introduced by the experimental procedures.
Blinding	Blinding was not meaningful as only routine in vitro assays for compound characterization have been performed.

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
☒	☐	Plants

Methods

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

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	HEK293T cells (CRL-1573) and CHO-K1 cells (CCL-61) were obtained from ATCC
Authentication	Cells were obtained from above-described sources and not further authenticated.
Mycoplasma contamination	Cell lines have been monitored/tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in the study.