nature research

Corresponding author(s): Andrew L. Goodman

Last updated by author(s): May 14, 2019

Reporting Summary

Nature Research 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 Research policies, see Authors & Referees 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	MassHunter Workstation Software (Version B.06.01, Agilent Technologies)
Data analysis	MassHunter Quantitative Analysis Software (Version 7.0, Agilent Technologies) MassHunter Qualitative Analysis Software (Version 7.0, Agilent Technologies) Mass Profiler Professional (Version 13.0, Agilent Technologies) Matlab (Version 2017b, MathWorks) Python (Version 3.6) trimmomatic-0.38 bioBakery tools: MetaPhlan2 v2.6.0, ShortBREAD v0.9.5, and KneadData v0.6.1 Qiime 1.8 USEARCH v11.0.667_i86linux32 RDKit (Open Source Chemoinformatics, http://www.rdkit.org) Cytoscape v3.4.0 Analysis pipeline scripts for analyzing data and generating figures are available on GitHub (https://github.com/mszimmermann/drug-bacteria-gene_mapping) and archived at Zenodo (https://doi.org/10.5281/zenodo.2827640).

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

Data are available from FigShare: https://doi.org/10.6084/m9.figshare.8119058.
All data generated during this study are included in this published article and Supplementary Tables 1-21.
Raw sequencing data were deposited on the publicly available ENA server (accession no. PRJEB31790).
Raw metabolomics data were deposited on the public repository MetaboLights (accession no. MTBLS896).

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	No prior assumptions were made regarding effect sizes and sample number were determined by technical and experimental feasibility.
Data exclusions	Tissue samples were excluded from analysis, if the sample weight was below 30mg. 30mg is twofold the maximal weight difference of the sample tubes used for weighing tissue samples before extraction. LC-MS data resulting from failed sample injection were excluded for further analysis.
Replication	Four to six biological replicates were performed for all experiments. The number of replicates for each specific experiment is indicated throughout the manuscript text and figure legends. All attempts of replication were successful.
Randomization	Combinatorial drug metabolism screen: drugs were randomly assigned to drug pools. Mouse experiments: mice were randomly assigned to study groups, while keeping genders equally distributed. Representative sequencing of gain-of-function library clones were selected randomly from the library plates.
Blinding	No blinding of experiments was required, as non of the experimental outcomes was based on (potentially biased) human judgment.

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
☐	☒	Animals and other organisms
☐	☒	Human research participants
☒	☐	Clinical data

Methods

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

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Germfree and conventional C57BL/6J, either gender (equally distributed between study groups), 9-16 weeks of age were used for this study.
Wild animals	The study did not involve wild animals.

Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	All experiments using mice were performed using protocols approved by the Yale University Institutional Animal Care and Use Committee.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	This project made use of previously collected, de-identified samples and no human research participants were recruited or studied during the project. We selected the right box above for "Human research participants" in order to clarify that that the samples studied during the project do NOT constitute a human research study as assessed by the Yale University Human Investigation Committee.
Recruitment	No human research participants were recruited for the project.
Ethics oversight	The use of previously collected, de-identified samples does not constitute a human study as assessed by the Yale University Human Investigation Committee.

Note that full information on the approval of the study protocol must also be provided in the manuscript.