nature research

Corresponding author(s): Ching Leang, Michael C. Jewett, Michael Köpke

Last updated by author(s): 11/1/2021

# 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 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<br>*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<br>*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 | Gene/Protein sequences were extracted using the JGI Integrated Microbial Genomes & Microbiomes (IMG/M) platform as described in the Online Methods. DNA sequence data was generated using MiSeq and built in MiSeq Control Software (MCS) (v3.1.0) as described in the Online Methods.<br><br>Metabolite data from HPLC was measured on Agilent HPLC 1100 and 1200 series systems and gas composition data was measured using gas chomatoraphy via Agilent 190 Micro GC systems. Data from both systems was collected using Agilent OpenLab CDS ChemStation (C.01.07) as described in the Online Methods<br><br>Proteomics LC-MS/MS data was collected on a Q Exactive Plus mass spectrometer running Xcalibur (v4.2.47; Thermo Scientific) data collection software. MS/MS data were searched using Proteome Discoverer (v2.3; Thermo Scientific) as described in the Online Methods.<br><br>Cell-free protein expression quantification data was generated using a MicroBeta2 instrument (Perkins Elmer) and built-in software.<br><br>No custom software was used for data collection. |
| Data analysis | Extracted and generated sequences were analyzed using Geneious (R9.1.8; Biomatters Ltd), ClustaW (v2.1), IQ-TREE (v1.6.12) as described in the Online Methods. DNA designs were generated using TeselaGen DESIGN module (v19) (TeselaGen Biotechnology; https://teselagen.com/).<br><br>Protein abundance data was log2 transformed, normalized, and QA/QC'd with InfernoRDN (v1.1.6687.24196). Statistical analyses of proteomics data were performed with Perseus (v1.6.14.0) as described in the Online Methods.<br><br>Metabolic modeling was performed using COBRApy (v0.8.2) and OptFlux (v3.2.10) with Gurobi Optimizer (v7.0.2), as described in the Online methods. |

Microsoft Excel was used for general data analysis. No custom software was used for data analysis.

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 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

All data presented in this manuscript are available as supplementary data files or specified databases below.

All genome sequences are available through the JGI Integrated Microbial Genomes & Microbiomes system (IMG/M) platform or NCBI Genbank under accession numbers provided in Supplementary Tab. 7.

All proteomics raw data is available at the ProteomeXchange Consortium via the MassIVE repository (ftp://massive.ucsd.edu/MSV000085940/) MassIVE accession: MSV000085940; ProteomeXchange accession: PXD020853.

Emission factors for LCA analysis are available and were taken directly from the GREET 1 model (https://greet.es.anl.gov/) and ecoinvent database (https://ecoinvent.org/the-ecoinvent-database/).

Any additional data may be available from the authors upon reasonable request.

# 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 | All sample sizes used are listed in the manuscript. No sample size calculation was performed. Sample size (n>=3) represents field standard and regular practice in literature. Quantification of cell-free expression levels to determine enzyme concentrations in cell-free acetone biosynthesis reactions was carried out as n = 2 as highly reproducible and small differences in levels would not impact conclusions. |
| Data exclusions | No data was excluded (i.e. no data was 'cherry-picked'). |
| Replication | All attempts at replication were successful and reproducibility is built in to the biological replicates measured in this study. All strains were run more than once (i.e. a separate biologically replicated set) and data was reproducible. Key findings were confirmed through a combination of experiments and models. |
| Randomization | Samples were organized by experimental variables then characterized fully and reported in completion. Therefore, samples were not randomized. |
| Blinding | Blinding was not relevant. Animal or human participants were not used in the study, a distribution of biolmolecules were measured that is unbiased and does not need blinds. |

# 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 |
| ☒ | ☐ Human research participants |
| ☒ | ☐ Clinical data |
| ☒ | ☐ Dual use research of concern |

## Methods

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