nature research

Corresponding author(s): Steven Singer

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

For enzymatic assays (Figure 1, 4, S1, S7): three technical replicates were chosen, which is standard in the field and provided acceptable SEM.

For mass spectrometry measurements for glycosylation (Table S11, Figure S11) one experiment was performed

2. Data exclusions

Describe any data exclusions.

None

3. Replication

Describe whether the experimental findings were reliably reproduced.

For protein purification (Figure S7), PAGE (Figure 3, 4c), proteomic identification (Table S9, S10; Figure S8, S9), glycosylation assay (Figure S10) and protease assay (Figure S13), experiments were performed 3-5 times (specified in text) and were nearly identical in all cases.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Not applicable

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Not applicable.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	A statement indicating how many times each experiment was replicated
☒	☐	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☒	☐	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☒	☐	The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	MaxBin 2.0- binning algorithm; Trimmoatics, read trimming; IDBA-UD, metagenome assembly; CheckM, bin analysis; MEGAN4, phylogenetic classification; Prodigal, gene prediction software; Bellerephon, chimera check; MEGA5, tree building software; MASCOT, proteomics analysis; Scaffold, proteomics analysis

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	No restrictions.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	Not applicable.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	Not applicable
b. Describe the method of cell line authentication used.	Not applicable
c. Report whether the cell lines were tested for mycoplasma contamination.	Not applicable
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	Not applicable

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	Not applicable

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	Not applicable