nature research

Corresponding author(s): Danna R Gifford

☐ 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.

Sample size was chosen to optimize detection of differences in proportion, such that e.g. the minimum difference that could be detected with power 0.8 was 0.07 for proportions at the extremes where the variance is lowest (e.g. or p=1 vs p=0.93), and 0.18 for proportions around p=0.5 where variance is highest (e.g. p=0.41, p=0.59)

2. Data exclusions

Describe any data exclusions.

To increase the breadth of phylogenetic coverage, genomes from the strain SBW25 were excluded due to its similarity with strain Pf0-1

3. Replication

Describe whether the experimental findings were reliably reproduced.

All attempts at replication were successful.

4. Randomization

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

Bacterial cells were distributed randomly from a larger population of cells to initiate selection lines.

5. Blinding

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

Control (i.e. no antibiotic) and treatment groups within each strain were transferred blinded.

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.	Statistics were performed in R with published open-source packages. Bioinformatics analysis was performed using a published pipeline incorporating published open-source tools. All details are described in the methods.

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).	N/A

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	N/A
b. Describe the method of cell line authentication used.	N/A
c. Report whether the cell lines were tested for mycoplasma contamination.	N/A
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.	N/A

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.	N/A

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.	N/A

nature research

Corresponding author(s): Dr. Danna Gifford

☐ Initial submission ☐ Revised version ☒ Final submission

Flow Cytometry Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

Data presentation

For all flow cytometry data, confirm that:

☒ 1. The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ 2. The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ 3. All plots are contour plots with outliers or pseudocolor plots.

☒ 4. A numerical value for number of cells or percentage (with statistics) is provided.

Methodological details

5. Describe the sample preparation.	Flow cytometry was performed on Accuri C6 (BD Biosciences, UK). The cell densities were adjusted to give around 1000 events per second.
6. Identify the instrument used for data collection.	BD Accuri C6 with CSampler attachment for 96 well plates
7. Describe the software used to collect and analyze the flow cytometry data.	BD Accuri C6 Software was used to collect the data. R flowCore and flowViz from Bioconductor were used to import, visualize and gate raw cytometry data.
8. Describe the abundance of the relevant cell populations within post-sort fractions.	Around 10000 events were acquired per sample. The gating procedure described below would typically preserve 7000-8000 events for subsequent analysis
9. Describe the gating strategy used.	During data acquisition, a lower cut off was set at 10,000 for FSC-H and at 8000 for SSC-H. In the gating pipeline, the events were automatically gated on size/shape by retaining the cells within 2 standard deviations around the median in the bivariate normal distribution of FSC-A and SSC-A ("norm2Filter" from flowCore package). Then, k-mean clustering algorithm was applied on fluorescence intensity FL1-H to differentiate fluorescent versus non-fluorescent cells ("kmeansFilter" from flowCore package). For each antibiotic concentration, we ensured that YFP-expressing strain can be well separated from non-fluorescent strains by overlaying non-mixed controls (overlap is usually less than 2% of the cells).

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information. ☒