

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 (<i>n</i>) 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 <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☒	☐ 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. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☒	☐ 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 <i>d</i> , Pearson's <i>r</i>), 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	Nanopore sequencing data was collected through commercial services of Plasmidsaurus. Pacbio sequencing data was collected through commercial services of Genewiz. SH800S software was used to collect the flow cytometry data. FlowJo (BD) was used to analyze the flow cytometry data. Sequencing analysis done using versions Python ≥ 3.8 Biopython ≥ 1.80 NumPy ≥ 1.20 Pandas ≥ 1.3 EMBOSS ≥ 6.6.0 Matplotlib ≥ 3.5 Seaborn ≥ 0.11 BLAST+ ≥ 2.12.0 NUPACK≥4.0. The data generated in this study was analyzed using custom Python scripts available on GitHub at: https://github.com/Metaxo/Sidewinder .
Data analysis	Data analysis details are provided in the Methods section where applicable.

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](#)

The sequencing data generated in this study have been deposited in the NCBI Sequence Read Archive with the accession number PRJNA1201800. Source data is available for this paper.

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

Reporting on race, ethnicity, or other socially relevant groupings

Population characteristics

Recruitment

Ethics oversight

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](https://www.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

Data exclusions

Replication

Randomization

Blinding

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

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A

Flow Cytometry

Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ 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).
- ☒ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	The fluorescent protein expressing E. coli clones were encapsulated in hydrogel microparticles and were suspended in phosphate-buffered saline (PBS) at pH 7.4 after culture and protein expression. The sample was then analyzed and sorted using FACS at a concentration of 10^6 particles per ml. Samples were stored at 4°C at all times.
Instrument	SONY SH800S Cell Sorter
Software	SH800S software was used to collect the flow cytometry data. FlowJo (BD) was used to analyze the flow cytometry data.
Cell population abundance	Approximately 500,000 hydrogels contained single colonies of bacteria expressing the fluorescent proteins designed using SideWinder. Sort purity for fluorescent positive colonies was greater than 95%.
Gating strategy	FSC and SSC were used to select the sample population and exclude noise or debris. Encapsulated clones were gated from the FSC/SSC plot using a threshold of 200,000 on each axis. The gated population was further analyzed using individual fluorescence channels, with positive populations identified based on fluorescence height-to-area plots. Positive populations were distinguished from the background by their increased fluorescence area and height.

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