nature portfolio

Corresponding author(s): Cyril Savin
Javier Pizarro-Cerda

Last updated by author(s): May 27, 2025

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 (*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	Human data collection was performed using Lagon 2.25.0 (Epiconcept, Paris, France) Published transcriptomes were downloaded from the Yersiniomics database (https://yersiniomics.pasteur.fr/) Y11 reference proteome was downloaded from NCBI website with RefSeq annotation from December 2022, accession GCF_000253175.1 (https://www.ncbi.nlm.nih.gov/) Y11 Gene Ontology terms were downloaded from the QuickGo website accessed on May 25th 2023 (https://www.ebi.ac.uk/QuickGO/)
Data analysis	BIGSdb 1.29 (Keith Jolley, Oxford University, UK) (https://pubmlst.org/software/bigsdb/) was used for genome-based taxonomic assignment. Variant calling was performed with Snippy v4.1.0. Putative recombinogenic regions were detected and masked with Gubbins v2.3.4. A maximum-likelihood phylogenetic tree was built with IQ-TREE2 v2.0.6. IQ-TREE2's ModelFinder was used to estimate the best variant of the GTR model. Tree visualization was created with iTOL v6. A pan-genome was constructed using PPanGGOLiN v1.1.72. Protein alignment was performed with MAFFT v7.407. IQ-TREE2 v1.6.7.2 was used for the Pan-genome phylogenetic analysis. Synapomorphies were analyzed using a dedicated script available at https://gitlab.pasteur.fr/GIPhy/findSynapomorphies/ Genes under positive selection pressure were detected with PAML v1.3.1. Bayesian evolutionary analysis to infer the evolutionary rates was performed using BEAST v1.10.4 Treetime version 0.8.6 was used to evaluate the presence of a temporal signal Structure analysis was realized on ChimeraX 1.9 GSEA was performed using the R package clusterProfiler v4.6.2

Correlations were calculated using the R package Hmisc v5.2.2 and displayed with the R package corrplot v0.95
We uploaded the R code on Zenodo : https://doi.org/10.5281/zenodo.15526016

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

Source data are provided with this paper. The genomic data generated in this study have been deposited in the European Nucleotide Archive database under accession code PRJEB19854. Sample ID starting with ERS and accession numbers starting with FWC: Ye.1 (ERS1580350, FWCF02000001-FWCF02000181), Ye.2 (ERS1580351, FWCE02000001-FWCE02000227), Ye.3 (ERS1580352, FWCB02000001-FWCB02000195), and Ye.4 (ERS1580353, FWCD02000001-FWCD02000173). https://www.ebi.ac.uk/ena/browser/view/PRJEB19854.
The mass spectrometry proteomics data generated in this study have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository70 under accession code PXD043567. https://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD043567.

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	Epidemiological data and bacterial isolates originate from a unique female patient as stated in the case description.
Reporting on race, ethnicity, or other socially relevant groupings	Non applicable
Population characteristics	Non applicable
Recruitment	Non applicable
Ethics oversight	The ethical committee of the University Hospital of Tours approved the study which is registered under the number 2025-032.

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

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Genomic evolution analysis was performed on all the strains (4) isolated from the patient. Phylogenetic reconstruction was performed using 259 additional strains.
Data exclusions	No data were excluded
Replication	Illumina NGS sequencing was performed one per sample with the exception of 3 strains whose genome were sequenced twice. All the replicates gave the same results. Proteomics experiment was performed once on 5 parallel bacterial growth for each strain. Growth curve were performed once with 1 technical replicate. Antibiotics susceptibility profiling was performed once wih 1 technical replicate. All survival assay upon antibiotics or cell exposure were performed 1 or 2 times with 2 to 6 replicated wells. Enzymatic assays were performed 2 times with 1 to 3 replicated wells.
Randomization	Phylogenetic reconstruction was performed using 259 additional strains randomly selected from strains isolated from 1982 to 2016.

Randomization	Processing of the proteomic samples was randomized after collection of bacterial culture. All other experiments were not randomized.
Blinding	Sample processing, library preparation, sequencing and taxonomic assignment were performed blinded. Subsequent bioinformatic analysis were not blinded as genomic differences between different strains were looked for. Processing of the proteomic samples was blinded during sample randomization. Antibiotics susceptibility profiling was blinded. All other experiments were not blinded.

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

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	ATCC
Authentication	Cell lines were authentified by ATCC, no further authentification was performed in our laboratory
Mycoplasma contamination	Cell lines tested negative for Mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	n/a

Plants

Seed stocks	Not applicable
Novel plant genotypes	Not applicable
Authentication	Not applicable