

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](#).
Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study.
For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

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	Clinical isolates and antimicrobial susceptibility data were generated during routine diagnostic and surveillance activities. Species were identified by MALDI-TOF mass spectrometry and whole-genome sequencing
Data analysis	Genome assembly: SPAdes v3.10.0. Genome annotation: Prokka v1.14.5. Pangenome analysis: PIRATE v1.0.5 and PANOPTICON. Sequence alignment: MAFFT. Phylogenetic inference: RAxML v8.2.11 and ClonalFrameML. AMR determinant detection: AMRFinderPlus. GWAS: pyseer. International-clone assignment: ICassigner. Machine-learning prediction: XGBoost implemented in BAMPS. Time-scaled phylogenetic analysis: BactDating and SkyGrowth.

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.

Custom analysis and figure-generation scripts are available at <https://github.com/Benizao1980/Acinetobacter-resistance-architectures>; reusable machine-learning code is available at <https://github.com/Benizao1980/BAMPS>.

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](#)

Genome data generated in this study are associated with NCBI BioProject PRJNA1518973. Isolate-level epidemiological metadata, population assignments, antimicrobial susceptibility measurements and accession identifiers are provided in Supplementary Table S1. Public genomes used for the global comparison are listed in Supplementary Table S2, and the independent validation collection is described in Supplementary Table S8. Complete pangenome, GWAS, AMRFinderPlus, machine-learning, global-prediction, temporal-analysis and figure-source datasets will be deposited in Figshare and linked by DOI. Genome assemblies and population annotations are additionally available through the pubMLST *Acinetobacter baumannii* database.

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	Patient sex was recorded in the source clinical metadata where available and was used only for descriptive epidemiological summaries.
Reporting on race, ethnicity, or other socially relevant groupings	Race and ethnicity information was not collected or analysed.
Population characteristics	The study included 284 anonymised clinical <i>A. baumannii</i> isolates collected between 2002 and 2018 from 69 hospitals in Russia, Belarus and Kazakhstan. Available metadata included specimen date and type, hospital, ward type, patient age and patient sex.
Recruitment	There was no prospective participant recruitment. Bacterial isolates were obtained during routine diagnostic and surveillance activities and were retrospectively analysed after anonymisation.
Ethics oversight	Clinical isolates were collected as part of routine diagnostic and surveillance activities. Samples and associated metadata were anonymised prior to analysis and no identifiable patient information was used. Ethical approval and consent requirements were waived in accordance with national regulations governing the use of bacterial isolates obtained through routine clinical care.

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	No prospective power calculation was performed as this was a retrospective genomic study using an available longitudinal clinical isolate collection.
Data exclusions	Assemblies inconsistent with <i>A. baumannii</i> or showing evidence of contamination, poor assembly quality or discordant metadata were excluded. Of 284 study isolates, 280 had suitable annotated assemblies for pangenome analyses and 276 had matched quantitative carbapenem MIC data for GWAS. Missing antimicrobial susceptibility measurements were handled on a per-antibiotic basis rather than by excluding the isolate from all analyses. Exclusion criteria were applied before downstream statistical analysis.
Replication	The study analysed independent clinical bacterial isolates rather than replicated experimental treatments.
Randomization	Randomisation was not applicable because this was an observational retrospective study with no experimental allocation of samples to treatment groups. Machine-learning training/test splits were generated computationally as described in the Methods.
Blinding	Blinding was not applicable because this was an observational genomic study and no experimental treatment allocation or subjective outcome assessment was performed.

Behavioural & social sciences study design

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

Study description	
Research sample	
Sampling strategy	
Data collection	
Timing	
Data exclusions	
Non-participation	
Randomization	

Ecological, evolutionary & environmental sciences study design

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

Study description	
Research sample	
Sampling strategy	
Data collection	
Timing and spatial scale	
Data exclusions	
Reproducibility	
Randomization	
Blinding	

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	
Location	
Access & import/export	
Disturbance	

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

Antibodies

Antibodies used	
Validation	

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	
Authentication	
Mycoplasma contamination	
Commonly misidentified lines (See ICLAC register)	

Palaeontology and Archaeology

Specimen provenance	
Specimen deposition	
Dating methods	
☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.	
Ethics oversight	

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	
Wild animals	
Reporting on sex	
Field-collected samples	
Ethics oversight	

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	
Study protocol	
Data collection	
Outcomes	

Dual use research of concern

Policy information about [dual use research of concern](#)

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes
☒	☐ Public health
☒	☐ National security
☒	☐ Crops and/or livestock
☒	☐ Ecosystems
☒	☐ Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes
☒	☐ Demonstrate how to render a vaccine ineffective
☒	☐ Confer resistance to therapeutically useful antibiotics or antiviral agents
☒	☐ Enhance the virulence of a pathogen or render a nonpathogen virulent
☒	☐ Increase transmissibility of a pathogen
☒	☐ Alter the host range of a pathogen
☒	☐ Enable evasion of diagnostic/detection modalities
☒	☐ Enable the weaponization of a biological agent or toxin
☒	☐ Any other potentially harmful combination of experiments and agents

Plants

Seed stocks	
Novel plant genotypes	
Authentication	

ChIP-seq

Data deposition

- ☐ Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- ☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links <i>May remain private before publication.</i>	
Files in database submission	
Genome browser session (e.g. UCSC)	

Methodology

Replicates	
Sequencing depth	
Antibodies	
Peak calling parameters	
Data quality	

Software

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

Instrument

Software

Cell population abundance

Gating strategy

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

Magnetic resonance imaging

Experimental design

Design type

Design specifications

Behavioral performance measures

Imaging type(s)

Field strength

Sequence & imaging parameters

Area of acquisition

Diffusion MRI

☐ Used☐ Not used

Preprocessing

Preprocessing software

Normalization

Normalization template

Noise and artifact removal

Volume censoring

Statistical modeling & inference

Model type and settings

Effect(s) tested

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference

(See [Eklund et al. 2016](#))

Correction

Models & analysis

n/a | Involved in the study

- ☐ ☐ Functional and/or effective connectivity
- ☐ ☐ Graph analysis
- ☐ ☒ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Graph analysis

Multivariate modeling and predictive analysis

XGBoost regression models were used to predict log2-transformed imipenem and meropenem MIC values from three genomic feature sets: AMRFinderPlus-derived AMR determinants, GWAS-derived locus-level features, and combined hybrid features. Models were evaluated using repeated training/test splits and a held-out internal test set, followed by independent external validation in 672 genomes. Performance was assessed using MAE, RMSE, R^2 , mean signed error and proportions of predictions within ± 1 and ± 2 two-fold MIC dilutions. Models were then applied without retraining to 1,588 global genomes. Complete feature definitions, model outputs and validation results are reported in Supplementary Tables S8-S10 and accompanying deposited data.

