nature portfolio

Corresponding author(s): Zhenglin Yang; Li Jiang; Hanbin Ma*

Last updated by author(s): Feb 20, 2026

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	Library quantification was performed using Qubit Fluorometer with Qubit 4 software. Fragment size distributions were analyzed on Agilent 4200 TapeStation using TapeStation Analysis Software. Sequencing was conducted on Illumina HiSeq 2500 with HiSeq Control Software (v2.2.58); raw BCL files were converted to FASTQ using bcl2fastq2 (Illumina).
Data analysis	Data analysis was performed using a standardized bioinformatics pipeline.The in-house Python script for RPM calculation is available from the corresponding author upon reasonable request. No other custom algorithms were used in this study.

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

Provide your data availability statement here.

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	n/a
Reporting on race, ethnicity, or other socially relevant groupings	n/a
Population characteristics	Residual fluid and tissue samples were collected from patients who underwent clinically indicated metagenomic next generation sequencing (mNGS) testing at the Department of Laboratory Medicine, Sichuan Provincial People's Hospital, School of Medicine, University of Electronic Science and Technology of China. Sample types included whole blood, bronchoalveolar lavage fluid (BALF), cerebrospinal fluid (CSF), ascites, synovial fluid, pus, urine, abscess drainage, biopsy tissue, and bone marrow.
Recruitment	Inclusion criteria were: (1) availability of sufficient sample volume (≥600 μL) after routine testing, and (2) complete accompanying clinical information. Exclusion criteria were: (1) insufficient sample volume, or (2) prior failure of the clinical mNGS assay. Written informed consent was obtained from all participants prior to sample collection.
Ethics oversight	The research protocol was approved by the Committee for Basic and Clinical Research Ethics of the Sichuan Academy of Medical Sciences & Sichuan Provincial People's Hospital (Approval No.: Ethics (Research) 2025 No. 557). All procedures performed were in accordance with the principles of the Helsinki Declaration, and this study complied with the ethical standards for research involving human participants in China.

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 statistical method was used to predetermine sample size. For analytical performance validation (e.g., LoD, reproducibility), sample sizes (n≥3) were chosen based on established standards in the field to ensure technical robustness. For clinical validation, the sample size (n=94) was determined by the availability of residual specimens meeting the inclusion criteria during the study period.
Data exclusions	Data exclusion criteria were pre-established. Clinical specimens were included only if they met the predefined criteria described in the Methods (sufficient volume, complete clinical info). No samples or data points were excluded from the analyses after data collection, except for the two discordant cases which are reported and discussed as false negatives.
Replication	All attempts at replication were successful. Analytical performance metrics (e.g., LoD, precision) were derived from at least three independent replicates (n≥3) as indicated in the figure legends. Inter-chip reproducibility was confirmed using two independent chip batches. The key findings of clinical concordance were validated using an independent cohort of 94 clinical specimens.
Randomization	For analytical performance experiments, samples were processed in a predefined order based on experimental conditions; randomization was not applicable. For the clinical validation, this was a retrospective study comparing two methods on the same set of specimens; therefore, allocation to experimental groups was not randomized.
Blinding	Investigators were not blinded to sample identity during on-chip library preparation due to the nature of the experimental setup. However, downstream bioinformatic analysis and pathogen identification were performed in a blinded manner, where the analyst was unaware of whether the sequencing data originated from AM-DMF or manual library preparation until the final comparative analysis was completed.

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

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	Not applicable. This study is a retrospective validation using residual clinical specimens and does not involve a prospective interventional clinical trial. Therefore, clinical trial registration was not required.
Study protocol	Not applicable, as this is not a prospective clinical trial. The study methodology, including specimen inclusion/exclusion criteria and analytical procedures, is fully described in the Methods section .
Data collection	Clinical specimens (whole blood, BALF, CSF, etc.) were collected between May and September 2025 from patients undergoing routine mNGS testing at the Department of Laboratory Medicine, Sichuan Provincial People's Hospital. All specimens were residual samples obtained after completing clinically indicated testing and were stored at -80°C until nucleic acid extraction. Data collection was performed retrospectively based on pre-established inclusion criteria (≥600 µL residual volume, complete clinical information).
Outcomes	The primary outcome was the diagnostic concordance between AM-DMF-Lib-prep and the conventional manual method in detecting pathogens from clinical specimens, assessed by positive percent agreement (PPA), negative percent agreement (NPA), and overall concordance. Secondary outcomes included quantitative correlation of pathogen reads (RPM) between methods, performance under low-input DNA conditions (3 ng, 5 ng), and assessment of library quality metrics (concentration, duplication rate, uniquely mapped reads).

Plants

Seed stocks	n/a
Novel plant genotypes	n/a
Authentication	n/a