nature research

Corresponding author(s): Hongzhe Li

Reporting Summary

Nature Research 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 Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	The paper uses both public data sets and simulated data sets. Links to access the public data are provided. DEMIC_simulator was used for generating the simulated data sets, and made available in a repository (see code availability).
Data analysis	The following software were used for data analysis: Bowtie 2 v2.3.2, samtools v0.1.19, MEGAHIT v1.1.1, MaxBin v2.2.4, MetaBAT v2.12.1, CheckM v1.0.7, PTRC v1.1 and iRep v1.10. Source code of DEMIC is available in a repository (see code availability).

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/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

The accession numbers and weblinks for all real data sets are provided in Methods. Simulated data are available on request from the corresponding author.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences

Study design

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

Sample size	This is a methodological paper reporting a computational method for estimating bacterial growth rates using shotgun metagenomic sequencing data. We used a synthetic dataset generated from 36, 36 and 50 sequencing data sets of Lactobacillus gasseri, Enterococcus faecalis and Citrobacter rodentium to evaluate the performances of our method (all available data sets of L. gasseri, E. faecalis generated previously were used). In the random tests, we also showed our method has stable performance when six or more samples were available. In the simulation tests, we generated 50 samples for totally 45 species from 15 genera of five phyla. For the real public data sets, all samples were used.
Data exclusions	No data were excluded.
Replication	The methods was evaluated for different groups of contig contamination, completeness and sample counts (See Randomization) by 10 times in each group. Finally, DEMIC was able to accurately estimate growth dynamics when 40% of contigs or more and when six samples or more were provided in these evaluations. The accuracy of DEMIC was not affected by contamination in the evaluations.
Randomization	To generalize our evaluation to diverse metagenomic data sets, three different types of random tests were performed to evaluate the effects of sample counts, fraction of contig contaminations and completeness of contig clusters on the performance. Specifically, groups of 3, 6, 10, 15, 20, 25 samples, groups with 5%, 10%, 15%, 20%, 25% and 30% of contig contaminations and groups of 30%, 40%, 50%, 60%, 70%, 80% and 90% completeness of contig clusters were considered. In the simulation tests, we randomly selected 15 genera with three species in each and generated a total of 50 samples for the 45 species. Average coverages and PTRs were randomly assigned for each species in each sample independently.
Blinding	Not relevant because this is a methodological study.

Materials & experimental systems

Policy information about availability of materials

n/a	Involved in the study
☒	☐ Unique materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Research animals
☒	☐ Human research participants

Method-specific reporting

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ Magnetic resonance imaging