

## 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

- |                                     |                                     |                                                                                                                                                                                                                                                            |
|-------------------------------------|-------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement                                                                                                                                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly                                                                                                                                    |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>                                                               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | A description of all covariates tested                                                                                                                                                                                                                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons                                                                                                                                        |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | 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) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings                                                                                                                                                           |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes                                                                                                                                     |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | 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

The Shell scripts used for viral identification and host prediction are available in the zenodo repository (<https://doi.org/10.5281/zenodo.17066837>). R codes used in this study are available in zenodo repository (<https://doi.org/10.5281/zenodo.18919645>)

Data analysis

metaWRAP v1.2.2; TrimGalore v0.6.4; MEGAHIT v1.2.9; dRep v3.4.5; GTDB-tk v2.3.2; VIBRANT v1.2.1; ViralVerify v1.1; PPR-Meta v1.1; Virsorter2 v2.2.4; DeepVirFinder v1.0; CheckV v1.0.1; CD-HIT v4.8.1; geNomad v1.5.2; BACPHLIP; vConTACT2 v0.11.3; minced v0.4.2; BLAST v2.9.0; ARAGORN v1.2.41; Bowtie2 v2.3.5.1; samtools v1.9; CoverM v0.6.1; Prodigal v2.6.3; eggno-mapper v2.1.12; DRAM-v v1.4.6; kofamscan v1.3.0; DefenseFinder v1.3.0; inStrain v1.9.0; R v4.2.2

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 viral genomes, protein clusters and host prediction results of Glacier-to-River Virome Catalogue (GtRVC) have been deposited in the zenodo repository (<https://>

doi.org/10.5281/zenodo.17066357). Raw sequence reads from four rivers used in this study have been deposited in the NCBI Sequence Read Archive under BioProject ID PRJNA1264540. Raw sequence reads from Tibetan Glacier Genome and Gene Catalogue (TG2G) were deposited in NCBI under BioProject ID PRJNA813429.

## 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 Not applicable.

Reporting on race, ethnicity, or other socially relevant groupings Not applicable.

Population characteristics Not applicable.

Recruitment Not applicable.

Ethics oversight Not applicable.

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)

## Ecological, evolutionary & environmental sciences study design

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

|                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Study description        | We sampled 597 metagenomes from the Yangtze, Yellow, Lancang, and Yarlung Tsangpo rivers and also collected 85 glacial metagenomes from public dataset (TG2G), and established the Glacier-to-River Virome Catalogue (GtRVC) to decode viral diversity, viral adaptation, host interaction and ecological influence across elevation gradients.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Research sample          | A total of 597 metagenomes were sampled from the Yangtze, Yellow, Lancang, and Yarlung Tsangpo rivers between 2014 and 2018, together with 85 additional metagenomes obtained from the public TG2G dataset.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| Sampling strategy        | The sampling sites were distributed across the Qinghai–Tibet Plateau sections of the four major rivers (Yangtze, Yellow, Lancang, and Yarlung Tsangpo) as well as their downstream reaches, covering diverse geological contexts. In addition, multiple types of samples—including water, particle, and sediment—were collected to explore microbial and biogeochemical differences across media. This spatially integrated design ensured representative coverage of the entire river systems and provided reliable samples for investigating viral diversity and ecological influence along the river flow.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Data collection          | At each sampling site, approximately 10 L of river water was collected from ~0.5 m below the water surface and transferred into two sterile polyethylene terephthalate (PET) bottles for downstream processing. All samples were transported in refrigerated conditions (~0–4°C) to local laboratories for further processing. For the Yellow River, Lancang River, and Yarlung Tsangpo River, the water samples were first filtered through 3.0 µm polycarbonate membranes (Millipore, USA) to capture particle-associated microbes. Then, the filtrate obtained from the 3.0 µm membrane filtration was filtered again through 0.22 µm polycarbonate membranes (Millipore, USA) to capture free-living microbes and cellular fraction viruses. Both the 3.0 µm and 0.22 µm size fractions were subjected to metagenomic sequencing, and the resulting samples were designated as “particle” and “water” types, respectively. As for the Yangtze River, the total planktonic community was directly collected by filtering through 0.22 µm polycarbonate membranes (Millipore, USA). All pre-treated filters were stored in 50mL sterilized centrifuge tubes. Surface sediment samples for each rivers were sampled at sites with a water depth of approximately 0.5 m, transferred into sterile 50-mL polypropylene tubes, and then kept on dry ice during transport to the laboratory. All samples were subsequently preserved at –80 °C until downstream DNA extraction. |
| Timing and spatial scale | A total of 682 metagenomic samples were included in this study, consisting of riverine samples collected from the Yangtze, Yellow, Lancang, and Yarlung Tsangpo rivers between 2014 and 2018, as well as glacial samples collected from the public dataset (TG2G). These samples are distributed across a latitudinal range of 21.67°–40.62° N and a longitudinal range of 75.32°–120.96° E, with elevations ranging from 0.7 to 6675 m, covering both plateau and non-plateau regions.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| Data exclusions          | All sequencing data were used for analyses without exclusion.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Reproducibility          | This study centers on bioinformatic analyses, all of which can be reproduced using publicly available software packages described in the Methods section.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Randomization            | The study was designed to discover viruses and their patterns from any possible condition. The work does not focus on the effect of environmental parameters                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |

Blinding

Blinding is not relevant to this study. This study is a massive descriptive discovery of previously unknown pattern of viruses along the glacier-to-river gradient, blinding treatment does not impact the findings.

Did the study involve field work?



Yes



No

## Field work, collection and transport

Field conditions

A total of 597 metagenomic samples were sampled from the Yangtze, Yellow, Lancang, and Yarlung Tsangpo rivers among 2014 and 2019, and 85 associated glacial samples from TG2G were also collected. The sampling sites spanned both the Qinghai-Tibet Plateau and the downstream regions, covering latitudes from 21.67° to 40.62° N, longitudes from 75.32° to 120.96° E, and elevations ranging from 0.7 m to 6675 m.

The headwater regions of the four rivers on the Qinghai-Tibet Plateau generally exhibited lower temperatures (mean annual temperature: -5 to 5 °C) and moderate precipitation (mean annual precipitation: 400–1000 mm), whereas the downstream regions experienced higher temperatures (mean annual temperature: 10 to 25 °C) and greater precipitation (mean annual precipitation: 600–2000 mm). This wide environmental gradient provides representative coverage of the transition from high-altitude plateau to lowland plain systems.

All samples were collected using sterile containers and transported under refrigerated or frozen conditions to preserve sample integrity before laboratory processing.

Location

The sampling sites are distributed across latitudes of 21.67°–40.62° N and longitudes of 75.32°–120.96° E, with elevations ranging from 0.7 m to 6675 m.

Access &amp; import/export

597 samples were collected within China and processed by the study team. All field sampling was conducted responsibly and in compliance with local, national, and international regulations. Sampling procedures followed the Chinese national standards GB/T 14581-1993, HJ 493-2009, HJ 494-2009, and HJ/T 91-2002, ensuring proper environmental protection and minimal habitat disturbance.

Disturbance

No disturbances were caused by this study.

## 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                                  |
|-------------------------------------|--------------------------------------------------------|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                 |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Plants                        |

### Methods

| n/a                                 | Involved in the study                           |
|-------------------------------------|-------------------------------------------------|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Plants

Seed stocks

Not applicable.

Novel plant genotypes

Not applicable.

Authentication

Not applicable.