nature portfolio

Corresponding author(s): Miranda de Graaf

Last updated by author(s): Sep 5, 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	no software was used
Data analysis	R 4.4.2, Python 3.8.20, FastP 0.23.4, CD-HIT, BWA-MEM 0.7.17, metaSPAdes 3.13.0, DIAMOND 2.1.10.16475, pyCoDaMath :https://bitbucket.org/genomicepidemiology/pycodamath/src/master/, tidygeocoder (version 1.0.5) Genomad (version 1.11.0), MAFFT 7.508, IQTREE 2.3.6. The metagenomic workflow used in this study is made available at https://github.com/EMC-Viroscience/Worp_etal_Global_Sewage_Virome.

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 raw sequencing datasets, metagenome assemblies and analysis results are available on the European Nucleotide Archive under project: PRJEB87273.

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	*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used.* *Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected.* *Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*
Reporting on race, ethnicity, or other socially relevant groupings	*Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status).* *Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.)* *Please provide details about how you controlled for confounding variables in your analyses.*
Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

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

Ecological, evolutionary & environmental sciences study design

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

Study description	Observational study of the viral genetic material in untreated global wastewater samples
Research sample	Untreated wastewater samples from 62 cities in 47 countries. DNA/RNA in each sample reflects the virome and enteric viruses present at a particular site and the (human) population that contributed to the samples through defecation
Sampling strategy	The study was widely advertised within and out of our network. No exact sample size calculation was performed, but we aimed to cover as many locations around the world as possible. The sample size was determined by voluntary cooperation of the WWTP. Each participant was instructed to sample ~1L untreated community wastewater before treatment using either sampling over 24H or 3 x 300 mL approach at least 5 minutes apart. The protocol and instructions are described in Hendriksen et al. (Nature Communications, 2019) in Supplementary Data 8.
Data collection	Collaborating partners partners were instructed on how to take the wastewater samples and how to prepare them for transport (frozen transport to Denmark DTU and later Erasmus MC). They were also instructed on how to fill out metadata forms submitted through Survey Monkey. Detailed instructions are described in Hendriksen et al. (Nature Communications, 2019). In brief, for each

	sampling event partners labeled two clean 1,000 mL plastic containers (no detergent/disinfectant residue) with country, city, collector, sample number, and date, sealing labels with tape; 1 L sewage was collected per container with headspace for freezing. Preferred sampling was a 24-h flow-proportional composite from the mid-stream at the WWTP inlet; where not feasible (e.g., untreated discharges), three ~300 mL grab samples ≥5 min apart were taken and pooled. The sewage temperature was recorded. Containers were closed tightly, wiped with alcohol, kept cold, and transported to the local laboratory within 8 h; each was bagged individually and stored at −80 °C as soon as possible. Claudia Schapendonk and Nathalie Worp were responsible for isolating TNA and sequencing and Nathalie Worp was responsible for gathering in silico data.
Timing and spatial scale	Biannual global campaigns were run to minimize hemispheric seasonality, with Winter (November/December) and Summer (June/July) sampling windows. Longitudinal sampling was performed in selected cities across all inhabited continents. All samples used for the viral analyses were collected from May 2017 through January 2019; outside these windows there was no routine sampling. Spatial scale: city level, sampling at WWTP inlets (typically one site per city). The study relied on convenience sampling with broad geographic recruitment, aiming for global coverage but ultimately determined by partner availability.Wastewater samples of 62 cities from 47 countries and 6 continents were obtained.
Data exclusions	No data were excluded from the analysis.
Reproducibility	In (Hendriksen et al, 2019 Nat.Communications) reproducibility analysis on number of samples taken 1 day apart was performed and use to validate the study design. In (Nieuwenhuijse et al. 2020 Nature Scientific Reports) data quality evaluation and technical factors that can influence analysis of viral data were performed. Furthermore, we also sequenced a technical replicate (same sample, repeated library/sequencing). The replicate showed consistent community profiles. All replication attempts were successful; none failed.
Randomization	Not relevant for the study; no treatment and control groups as are commonly used in clinical/lab studies.
Blinding	Blinding is not applicable as this study does not involve treatment or control groups, but rather an observational analysis of wastewater viromes across different locations and time points

Did the study involve field work? ☐ Yes ☒ No

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

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*