RNA-Seq of untreated wastewater to assess COVID-19 and emerging and endemic viruses for public health surveillance

Stephen R. Stockdale,[a] Adam M. Blanchard,[b] Amit Nayak,[c] Aliabbas Husain,[c] Rupam Nashine,[c] Hemanshi Dudani,[c] C. Patrick McClure,[d,e] Alexander W. Tarr,[d,e,f] Aditi Nag,[g] Ekta Meena,[g] Vikky Sinha,[g] Sandeep K. Shrivastava,[h] Colin Hill,[a] Andrew C. Singer,[i] Rachel L. Gomes,[j] Edward Acheampong,[j,k] Saravana B. Chidambaram,[l] Tarun Bhatnagar,[m] Umashankar Vetrivel,[n,o] Sudipti Arora,[g,**] Rajpal Singh Kashyap,[c,***] and Tanya M. Monaghan[d,p,*]

[a]APC Microbiome Ireland, University College Cork, Co. Cork, Ireland
[b]School of Veterinary Medicine and Science, University of Nottingham, Nottingham, United Kingdom
[c]Research Centre, Dr G.M. Taori Central India Institute of Medical Sciences (CIIMS), Nagpur, Maharashtra, India
[d]National Institute for Health Research Nottingham Biomedical Research Centre, Nottingham University Hospitals National Health Service Trust, Nottingham, United Kingdom
[e]Wolfson Centre for Global Virus Research, University of Nottingham, Nottingham, United Kingdom
[f]Queen's Medical Centre, School of Life Sciences, University of Nottingham, Nottingham, United Kingdom
[g]Dr. B. Lal Institute of Biotechnology, 6-E, Malviya Industrial Area, Malviya Nagar, Jaipur, India
[h]Centre for Innovation, Research & Development, Dr. B. Lal Clinical Laboratory Pvt. Ltd., Malviya Industrial Area, Malviya Nagar, Jaipur, India
[i]UK Centre for Ecology and Hydrology, Wallingford, United Kingdom
[j]Food Water Waste Research Group, Faculty of Engineering, University of Nottingham, United Kingdom
[k]Department of Statistics and Actuarial Science, University of Ghana, P.O. Box, LG 115, Legon, Ghana
[l]Department of Pharmacology, JSS College of Pharmacy, JSS Academy of Higher Education & Research, Mysuru, 570015, KA, India
[m]ICMR-National Institute of Epidemiology, Chennai, Tamil Nadu, India
[n]National Institute of Traditional Medicine, Indian Council of Medical Research, Belagavi, 590010, India
[o]Virology and Biotechnology Division, ICMR-National Institute for Research in Tuberculosis, Chennai, 600031, India
[p]Nottingham Digestive Diseases Centre, School of Medicine, University of Nottingham, Nottingham, United Kingdom

The Lancet Regional Health - Southeast Asia 2023;14: 100205

Published Online 8 May 2023
https://doi.org/10.1016/j.lansea.2023.100205

Summary

Background The COVID-19 pandemic showcased the power of genomic sequencing to tackle the emergence and spread of infectious diseases. However, metagenomic sequencing of total microbial RNAs in wastewater has the potential to assess multiple infectious diseases simultaneously and has yet to be explored.

Methods A retrospective RNA-Seq epidemiological survey of 140 untreated composite wastewater samples was performed across urban (n = 112) and rural (n = 28) areas of Nagpur, Central India. Composite wastewater samples were prepared by pooling 422 individual grab samples collected prospectively from sewer lines of urban municipality zones and open drains of rural areas from 3rd February to 3rd April 2021, during the second COVID-19 wave in India. Samples were pre-processed and total RNA was extracted prior to genomic sequencing.

Findings This is the first study that has utilised culture and/or probe-independent unbiased RNA-Seq to examine Indian wastewater samples. Our findings reveal the detection of zoonotic viruses including chikungunya, Jingmen tick and rabies viruses, which have not previously been reported in wastewater. SARS-CoV-2 was detectable in 83 locations (59%), with stark abundance variations observed between sampling sites. Hepatitis C virus was the most frequently detected infectious virus, identified in 113 locations and co-occurring 77 times with SARS-CoV-2; and both were more abundantly detected in rural areas than urban zones. Concurrent identification of segmented virus genomic fragments of influenza A virus, norovirus, and rotavirus was observed. Geographical differences were also observed for astrovirus, saffold virus, husavirus, and aichi virus that were more prevalent in urban samples, while the zoonotic viruses chikungunya and rabies, were more abundant in rural environments.

Interpretation RNA-Seq can effectively detect multiple infectious diseases simultaneously, facilitating geographical and epidemiological surveys of endemic viruses that could help direct healthcare interventions against emergent and

*Corresponding author. National Institute for Health Research Nottingham Biomedical Research Centre, Nottingham University Hospitals National Health Service Trust, Nottingham, United Kingdom.
**Corresponding author.
***Corresponding author.
E-mail addresses: tanya.monaghan@nottingham.ac.uk (T.M. Monaghan), sudiptiarora@gmail.com (S. Arora), rajpalsingh.kashyap@gmail.com (R.S. Kashyap).

pre-existent infectious diseases as well as cost-effectively and qualitatively characterising the health status of the population over time.

Funding UK Research and Innovation (UKRI) Global Challenges Research Fund (GCRF) grant number H54810, as supported by Research England.

Keywords: COVID-19; Endemic viruses; RNA-Seq; SARS-CoV-2; Sewage surveillance; Wastewater-based epidemiology

Research in context

Evidence before this study

High-throughput RNA-sequencing (RNA-Seq) is an emerging tool to monitor, identify, and explore the diversity of viral pathogens circulating among human and livestock populations. However, to date, few studies have implemented this technique to study wastewater matrices for the purpose of understanding human population health. Wastewater-based epidemiology is a promising strategy to detect pathogens such as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and may serve as an early warning system for disease outbreaks, and also as a routine monitoring system for public health departments to understand the state of health of their nation. A search of primary research articles published in PubMed up until June 18, 2022, using the search terms "RNA-Seq" and "wastewater" or "RNA-Seq" and "sewage" or "wastewater virome" returned 54, 19, and 38 studies, respectively describing transcriptome sequencing findings, of which the majority were biased towards specific microorganisms and were mainly conducted in high-income countries. Two studies detected potential human pathogenic viruses in wastewaters from Latin America and China. Additionally, three studies employed RNA-Seq approach for the detection and mutational/variant analysis of SARS-CoV-2 in wastewater.

Added value of this study

To our knowledge, this study is the first to use unbiased RNA-Seq to obtain a snapshot of the RNA virosphere in untreated wastewater from India. This observational wastewater-based proof-of-principle epidemiological study conducted during the second wave of the COVID-19 pandemic, revealed the detection of gastrointestinal, parenterally transmitted, and zoonotic viruses. Of note, our findings reveal for the first time the detection of endemic zoonotic viruses in wastewater samples including Jingmen tick virus, rabies virus and chikungunya virus. SARS-CoV-2 was detected in 59% of sampling locations in Nagpur district, Central India, showing a rural area predominance. Hepatitis C virus was the most frequently detected virus of which the most prominent genotype was genotype 2, identified in 113 locations and co-occurring 77 times with SARS-CoV-2. Sequencing analysis also revealed the presence of influenza A virus, norovirus, and rotavirus in addition to several other diarrhoeal-causing viruses including astroviruses, saffold virus, husavirus and aichi viruses.

Implications of all the available evidence

This study supports the use of wastewater RNA-Seq as a valuable qualitative tool for public health surveillance and can be implemented for the identification of endemic and problematic viruses. RNA-Seq as applied to wastewater-based epidemiology may also assist in the design and implementation of national vaccination programs against emergent and pre-existent infectious diseases.

Introduction

Successful detection of SARS-CoV-2 and other pathogenic microbes underscores the potential utility of wastewater-based epidemiology (WBE) for community surveillance of infectious diseases. Wastewater surveillance can complement clinical diagnostic surveillance by providing early signs of potential transmission to enable more active public health responses.[1] However, many WBE studies at the outset of the pandemic focused their efforts on wastewater treatment plants for WBE, as this offered a cost-effective insight into large populations. However, over time, interest grew in understanding sub-catchment scale epidemiological trends, leading to increasing numbers of studies focused on within sewer network and near-source WBE.[2] Most studies describing wastewater surveillance have been conducted in high-income countries with well-developed sewage infrastructure.[3] Nevertheless, many households in low-to middle income countries (LMIC) are not connected to sewage networks due to poor sanitation infrastructure, with consequent mismanagement of sewage. As such, surveillance approaches that commence with sampling WWTP influent in LMICs might not be representative of actual disease prevalence in the wider community within the city under surveillance. Few studies have reported SARS-CoV-2 wastewater surveillance data in

low sanitation settings such as India,[4–9] emphasising that more investigation of wastewater surveillance approaches is urgently warranted in LMIC settings who at present might be deprived of the early-warning benefits of WBE.[10] Another important limitation is the lack of sequencing studies on wastewater samples, particularly high-throughput RNA sequencing (RNA-Seq) analysis.[10–18] This is in contrast to DNA sequencing studies such as Singh et al. who recently reported on the bacterial composition of the Indian sewage microbiome using 16 S rRNA gene amplicon sequencing.[19]

In view of these limiting factors, our primary aims were to firstly explore the feasibility of undertaking a wastewater-based RNA-Seq study of untreated wastewater samples in a resource limited and environmentally challenging LMIC setting, here Nagpur district, Maharashtra, Central India. Our secondary aims were to showcase the power of genomic sequencing to assess the abundance of multiple viral pathogens simultaneously, to examine the co-occurrence of SARS-CoV-2 with other viruses, and to compare the geographic distribution of viruses in wastewater samples from urban and rural areas in wastewater samples of this selected region. We thus concentrated our sampling efforts during the second wave of the COVID-19 pandemic in Nagpur, India. Nagpur was the epicentre of the second deadly wave of COVID-19 in India, which was driven by the SARS-CoV-2 Delta variant B.1.617 and saw daily peaks ranging from 5000 to 7000 cases during the month of April 2021 (figures from Nagpur Municipal Corporation).

Methods

Approvals for wastewater sample collection

As this was an environmental wastewater sampling study, no formal ethics was required from the respective institutions. However, official permissions for sample collection were taken from Nagpur Municipal corporation (NMC).

Study design and sample collection

We performed a retrospective cross-sectional metatranscriptional analysis of untreated wastewater samples in Nagpur district, Maharashtra, India from 3rd February to 3rd April 2021 during the second wave of the COVID-19 pandemic. For wastewater sampling, 422 grab samples of volume 1 L each were collected in sterile plastic leak proof containers from the main sewer drainage lines of 10 different urban municipality zones and from open drains/surface water resources in rural areas from Nagpur district. Typically, wastewater samples were collected by directly filling sample containers up to the 1 L mark. After sample collection, the sample containers were sanitized with 70% ethanol, labelled with zonal information and sealed in zip-lock bags. One member of the field team created a separate performa sheet for each grab sample and filed it while the other team members collected the samples simultaneously. All samples were transported to the CIIMS Research department under cold chain while being kept at the ideal temperature (2–4 °C) in a thermocol box with ice gel packs. All discrete individual samples were then pooled together in equal volume, spanning geography and time (3–4 random samples) under aseptic conditions. Based on the above sampling strategy, a total of 140 composite samples were prepared which included 112 composites prepared from urban areas and 28 from rural areas of Nagpur district. All sampling was conducted in the morning hours between 07:30 and 12:00 when flow rate was presumed to be highest due to an anticipated higher defecation frequency in the morning as previously reported by Heaton et al.[20] maximizing the likelihood of detecting viruses. Area demographics, in addition to GPS co-ordinates, temperature, relative humidity and rainfall were also recorded during the sampling period. All procedures of sample collection were done under biosafety conditions as per the standard operating procedures recommended by the Government of India.[21] All samples were labelled according to zonal information, sealed in zip lock pouches, and transported under cold chain (4± 2 °C) to the Research Department, Central India Institute of Medical Sciences (CIIMS). Samples were then transferred to the Environmental Biotechnology Laboratory at Dr. B. Lal Institute of Biotechnology, Jaipur under cold chain transportation for pre-processing of wastewater and RNA extraction for sequencing as previously described (Fig. 1).[8,22]

Storage and pre-processing

Upon receipt of all wastewater samples at the Dr. B. Lal Institute of Biotechnology, each sample was checked to ensure that cold chain transport was maintained during transportation. Wastewater samples were kept at 4 °C until they were pre-processed within 24 h for nucleic acid extraction. The sample containers were first externally sterilised for 30 min using UV treatment followed by heat inactivation done at 70 °C for 90 min in a sonicating water bath.[8] Samples were brought to room temperature and the grit particles were removed by crude filtration using Whatmann paper and then filtered using a vacuum filter assembly via a 0.45 μm membrane. Each sample's filtrate was placed in a new 50 mL falcon containing 0.9 g sodium chloride (NaCl) and 4 g polyethylene glycol (PEG). The contents were dissolved by gentle manual mixing. Samples were then centrifuged at 4 °C for 30 min at 5750×*g*. The pellet was re-suspended in the RNA/DNA shield solution provided in the ZYMOBIOMICS™ kit.

Nucleic acid extraction

Sample pellets were re-suspended in RNA/DNA shield buffer and were then further processed for manual extraction of nucleic acid by the ZymoBIOMICS™ kit

Fig. 1: **Schematic outline of study methodology.**

(catalogue no. R2136) according to the manufacturer's instructions. Samples were then vortexed mixed for 15 min and centrifuged at 10,000×*g* for 1 min at and 20 µL of proteinase K was added to the sample and incubated at room temperature for 30 min. A volume of 500 µL of DNA/RNA lysis buffer was next added to the 200 µL sample and mixed well. This was followed by a DNA separation step by adding 30 µL of Mag-Binding beads in the above mixture and mixing well for 20 min. The microfuge tubes were then transferred to the magnetic stand until all the beads pelleted. Clear supernatant containing RNA was transferred to a fresh tube.

RNA purification

For RNA purification, 700 µL of 95–100% ethanol was added to the supernatant and mixed well. Next, 30 µL of MagBinding bead was added to the sample and mixed

well for 10 min. The microfuge tubes were then transferred to the magnetic stand until the beads pelleted. Next, 500 µL of MagBead DNA/RNA wash was added, and the sample was mixed, the beads were pelleted and supernatant discarded. This wash step was repeated. Next, 500 µL of 95–100% ethanol was added, the beads pelleted, and supernatant discarded. This wash step was repeated once. The pellet was then air dried until matt. Finally, 50 µL of ZymoBIOMICS™ DNase/RNase-free water was added and the sample mixed well for 5 min. The supernatant was transferred into a fresh tube and this latter step repeated once more. The supernatant was stored at –70 °C. The quality and quantity of the RNA sample was determined by using Eppendorf biospectrometer®.

Library preparation

Extracted RNA was prepared for sequencing using the Illumina TruSeq Total RNA Library Preparation Kit (Illumina, USA). RNA was reverse transcribed using a random hexamer and superscript II transcriptase (Invitrogen, USA) and a second strand synthesised. The ds cDNA was purified using AmpureXP beads (Beckman & Coulter, USA) before fragmentation to approximately 300bp using a Covaris M220 (Covaris Inc, USA). The fragmented ds cDNA was then end repaired, A-tailed and had Illumina adaptors ligated before size selection and further purification using AMPureXP beads (Beckman & Coulter USA). The ds cDNA was then PCR amplified to incorporate index adaptors (98 °C for 30 s, followed by 15 cycles of 98 °C for 10 s, 60 °C for 30 s, 72 °C for 30 s and a final extension of 72 °C for 5 min). The PCR product was purified again using AMPureXP beads (Beckman & Coulter USA) before being assessed on an Agilent Tapestation using the high sensitivity D1000 tape (Agilent USA). Finally, the libraries were quantified using a Qubit BR DNA assay (Invitrogen USA) and then normalised and pooled according to manufacturer recommendations.

Sequencing

Sequencing was performed on an Illumina HiSeq 2500, employing paired end 2 × 150bp chemistry to generate approximately 5 GB of data per sample. Reads were provided by Eurofins Genomics Bengaluru, India as quality-checked and adaptor-free. Compressed reads were downloaded individually from Eurofins India's cloud-based ShareFile infrastructure. The quality of RNA-Seq reads were further assessed using Fastqc (v0.11.5), with an additionally more stringent quality-based processing performed using Trimmomatic (v0.36; SLIDINGWINDOW:4:30 LEADING:3 TRAILING:3 MINLEN:50). Potential human reads were removed from the RNA-Seq data using Kraken 2 (v2.0.8), after the kraken2-build command was utilised to ftp download and compile a *Homo sapien* (Hsa) database. Finally, paired end read files were repaired using BBMap's "repair.sh" script.

Virome analysis

A custom viral database of human-infecting RNA viruses was built (March 7th, 2022) using the following Boolean search in NCBI's web portal: "Viruses [Organism] AND txid2559587 [Organism:exp] AND srcdb_refseq [PROP] NOT wgs [PROP] NOT cellular organisms [ORGN] NOT AC_000001:AC_999,999 [PACC] AND ("vhost human" [Filter])". Processed RNA-Seq reads were aligned to the RNA viral database using Bowtie 2 (v2.3.5.1) in end-to-end mode. The number of reads aligned, and their breadth of coverage, per viral sequence was calculated using SAMTools (v1.10) and BEDTools (v2.18). However, a BED (Browser Extensible Data) formatted version of the human-infecting RNA viral database was necessary and created using the following commands: (i) samtools faidx [input fasta file]; and (ii) awk 'BEGIN {FS = "\t"}; {print $1 FS "0" FS $2}' [fai indexed input fasta file] > [output bed format file]. A "grep '@' –c" command was used to calculate the overall number of paired-end and unpaired sequencing reads per sample after concatenation (median: 30,677,058; SD: 13,271,310).

Read mapping summaries were imported into R Studio v3.6.1 for analysis. Dataframes and matrices were manipulated as necessary using the reshape 2 package. Filtering criteria of (i) ten or more RNA-Seq reads mapped per viral genome and (ii) a breadth of coverage of 5% or more was applied. A negative control buffer-only sequenced sample was used to normalise the read alignment counts per sample. Subsequently, viral abundances were normalised by transforming the alignment counts to reads per kilobase per million reads (RPKM). The multiple raw metadata records of mixed composite samples were used to generate a single representative metadata entry per sample. Where metadata variables were numeric, the median value was chosen, although for character variables, either the most frequent or middle value after sorting was chosen (for further details, see Supplementary Tables S1–S3). For natural composite samples lacking locality metadata, the following three criteria were applied sequentially to populate missing values: (1) samples with a longitude ≥79.5 were deemed as a village locality, (2) samples with a latitude ≥21.1 were assigned to an urban locality, and (3) the remaining unknown values with a latitude ≤21.1 were considered peri-urban.

Manuscript images were produced using the ggplot2 package with the ggpubr extension. Colours were obtained from the R package pals. The representative map of the Nagpur district was obtained from Google maps, with the image cropped using Cartesian points overlapping with the longitudinal and latitude collection locations of untreated sewage samples. Clustering of co-occurring viruses was performed using base R packages as follows. The RPKM normalised abundances of genomic segments and genotypes of related viruses were aggregated after simplifying each virus' name.

Viruses within samples with an RPKM value greater than zero were deemed present. The measure of similarity and dissimilarity between samples was calculated using Euclidean distances, with subsequent hierarchical clustering using the Ward.D2 method. All images were arranged into their final multi-plot figures using Inkscape (v1.1).

Statistical methods

Median comparisons between two groups, or for three or more groups, were performed using the Wilcoxon or Kruskal–Wallis tests, respectively, with Bonferroni false discovery rate correction. Normality of data was assess using the Shapiro–Wilks test. As a result, Spearman rank-order correlations were performed. Statistical differences between two categorical variables were determined using the Chi-squared test, except for small group comparisons that were performed using Fisher's exact test. Statistical significance between three or more categorical variables was established using ANOVA tests. All statistical tests were carried out at significance level of $P < 0.05$.

Role of funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

140 untreated wastewater samples were collected during a COVID-19 wave in the Nagpur district of Central India (February–April 2021). Of the 140 samples, 83 samples were collected from urban locations where sewage from multiple sewerage systems lines were pooled, termed "pooled piped sewage" samples. There were an additional 29 urban and 28 rural "mixed composite" samples that were generated by pooling geographically similar samples from the same period. This was performed to ensure all wastewater samples were heterogeneous for the catchment areas they represent. Given the study's sampling size restrictions, and the nature by which samples were collected, the metadata associated with each sample was derived from the median of numeric variables or most frequently coded character variable. Therefore, nuanced and potentially interesting data may be overlooked. For instance, where a market or hospital is in the vicinity of one of the untreated wastewater samples, these were pooled with other samples from a similar locality. Therefore, additional investigations using this study's data as a resource should note Supplementary Tables S1–S3.

RNA-Seq reads from all 140 samples were assessed against a RefSeq-derived custom database of all known human-infecting RNA viruses (updated March 2022; see Methods). Of the 822 sequences composing this viral database, which encompasses viruses with both monopartite and segmented genomes and viral genotypes, 63 unique sequences were detected at least once.

Non-gastrointestinal human viruses are identifiable in untreated wastewaters

SARS-CoV-2 was detectable in 83 of the 140 untreated wastewater samples (Fig. 2a). Furthermore, SARS-CoV-2 accounted for >1% of the total metagenomic reads for eight samples and even recruited 35.3% of the RNA-Seq reads for the rural mixed composite sample, CR4. Although there was an uneven sampling of urban and rural sites (Chi-Squared test, $P < 0.001$), there was a statistically significant difference in the normalised abundance of samples positive for SARS-CoV-2 by geography. Rural samples contained more SARS-CoV-2 than urban samples (Bonferroni adjusted Wilcoxon test, $P = 0.014$). This statistical significance was lost when SARS-CoV-2 was assessed by more stratified localities (i.e., urban, urban/peri-urban, peri-urban, slum, and village). The reads per kilobase per million reads (RPKM) abundances of SARS-CoV-2 were also assessed against temperature and humidity. Rank-order Spearman correlations showed no significant relationship between viral normalised abundance and temperature (rho = 0.15, $P = 0.085$), whereas a significant correlation between SARS-CoV-2 and humidity was observed (rho = –0.29, $P < 0.0001$).

Beyond SARS-CoV-2, untreated wastewater revealed an abundance of co-occurring human-infecting viruses (Fig. 2b). This included gastrointestinal viruses, parenterally transmitted viruses (i.e., outside of the intestines), and zoonotic viruses. However, when viral co-occurrence was arranged by hierarchical clustering, no clear distinction was observed between sampling locality and the viruses detected. Hepatitis C virus (HCV) was the most frequently detected virus in the untreated wastewater samples, present in 113 of the 140 samples (Fig. 3a). HCV and SARS-CoV-2 co-occurred in 77 wastewater samples, with only six samples positive for SARS-CoV-2 without HCV. The most prominent HCV genotype detected in this study was genotype 2 (n = 113), followed by genotype 1 (n = 65) and genotype 6 (n = 32). All HCV genotype 6 positive samples also contained genotype 1, and again, all genotype 1 positive samples were contained within the genotype 2 positive samples. Therefore, while HCV genotype sequences can differ by as much as 33% over their genome[23] conserved HCV regions may contribute to metagenomic-derived misclassifications of HCV genotypes.

Influenza A virus is continuously monitored for its potential to cause a pandemic-scale outbreak, with frequent mutations and genomic segment rearrangements, between human and animal-infecting strains, altering its infectious and immunogenic properties. In comparison to SARS-CoV-2, influenza A was infrequently detected in the wastewater samples analysed in

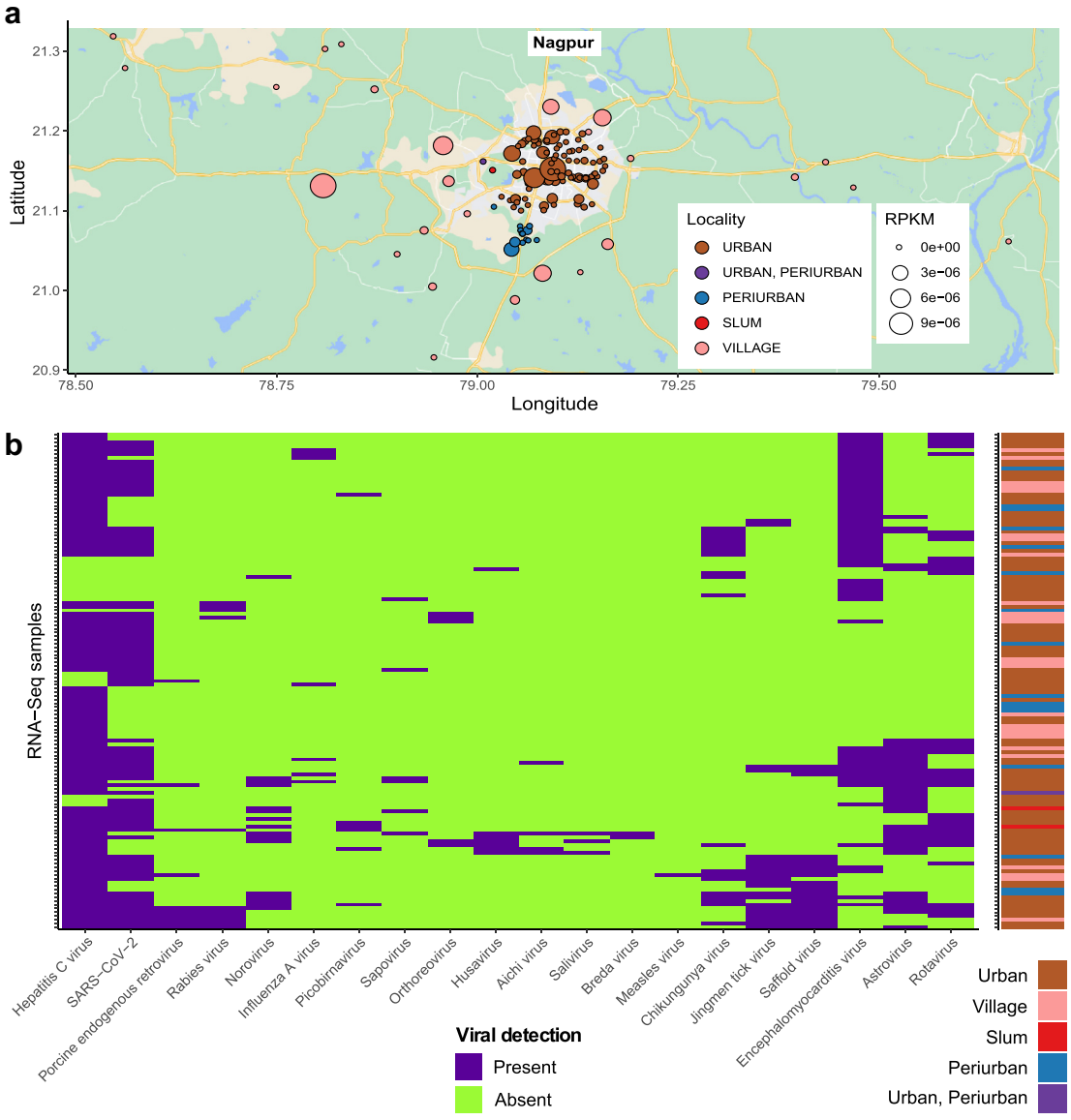

Fig. 2: **RNA-Seq to assess emerging and pre-existing pathogenic viruses in untreated sewage samples. (a)** The normalised abundance and best-approximation of geographical position of SARS-CoV-2 detected in untreated sewage samples from the Nagpur district of Central India. **(b)** The presence/absence co-occurrence of endemic viruses from the Nagpur district of Central India. RPKM, reads per kilobase per million reads.

this study (Fig. 3b). However, the rural mixed composite sample, CR26, harboured influenza A (H1N1) genomic segments 1, 2, 3, and 7. Whereas the urban mixed composite sample, C54, contained influenza A (H1N1) genomic segments 1, 2, and 5. C54 was recorded as an urban wastewater sample i.e., sampling location with a developed sewage infrastructure and CR26 was a rural wastewater sample collected from open drains, having an underdeveloped or poor sanitation system. It was also observed that in rural locations, there was a population of livestock nearby, which could possibly contribute to the finding.

RNA-Seq can detect faecal-oral viruses that disproportionally affect children

Rotaviruses A (RVA) and C (RVC) were detected in our untreated wastewater samples (Fig. 4a), with the magnitude of detection potentially inflated by a seasonal increase associated with this disease.[24] Urban samples C77, C94, C73, and C22, simultaneously contained nine, seven, five, and four rotavirus (RV) genomic segments, respectively. Whereas an additional three, nine and twenty samples contained three, two, or only one RV genomic segment, respectively. The most frequently detected rotaviral genomic segments were RVA segment

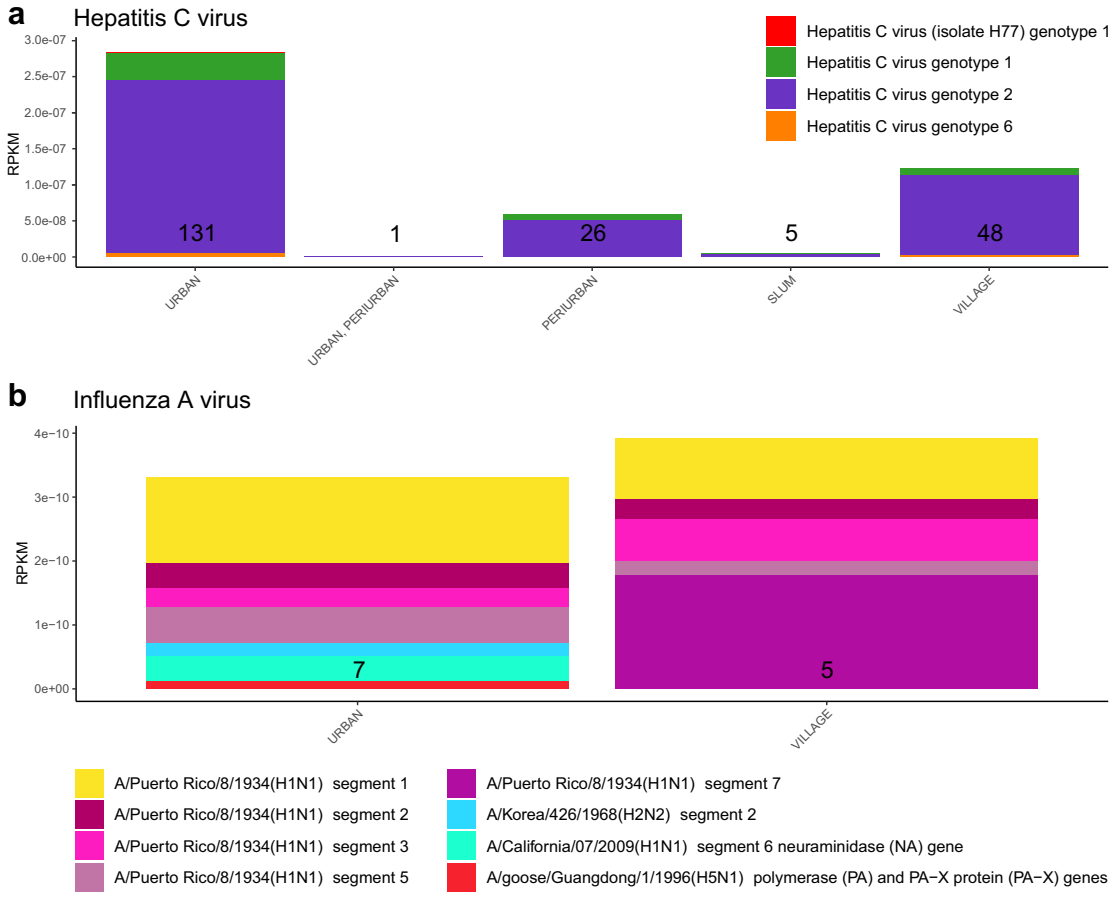

***Fig. 3:* Abundance and distribution of Hepatitis C virus and Influenza A virus in untreated sewage samples.** The normalised abundances of putative **(a)** Hepatitis C viral genotypes and **(b)** Influenza A viral genomic segments identified in untreated sewage samples. The number of samples with positive RPKM values that were accumulated per location are indicated at the base of each bar. RPKM, reads per kilobase per million reads.

11 and segment 6, present in 31 and 10 samples, respectively. RVA segment 6 co-occurred with segment 11 seven times. RVC genomic segments were detected more sporadically in untreated sewage, with no specific RVC segment present in more than 3 of the 140 samples, indicative of the biased urban sampling in this study and less frequent association of RVC with humans. Recalling that data manipulation was necessary to generate representative metadata for mixed composite samples, there were three sampling locations observed as having underdeveloped sewage infrastructure and four areas predominantly composed of open drains.

In our analysis of untreated wastewater, norovirus was only detected in urban localities (Fig. 4b). The Norovirus genogroup GII was the most frequently detected and abundant genogroup observed. The putative detection of norovirus genogroups GI and GIV both only occurred in single samples, whereas samples C94, C22, and C88, contained five, four, and three noroviral genomic segments, respectively. Where metadata was available for norovirus positive mixed composite samples, one and six locations were recorded as open drains with a properly developed sanitation infrastructure, and closed drains and with underdeveloped sanitation, respectively. Finally, as sample C22 was positive for both rotavirus and norovirus, we further investigated the metadata associated with this sample. Searching the longitude and latitude coordinates of C22 revealed it is centrally located and within approximately 2 km of five hospitals and primary health centres, which may reflect the emergence of nosocomial infection.

A multitude of human pathogenic viruses can be identified in untreated wastewater through RNA-Seq

Several additional diarrheal-causing viruses were identified in the untreated wastewater samples from the Nagpur district. In our analysis, astroviruses and saffold

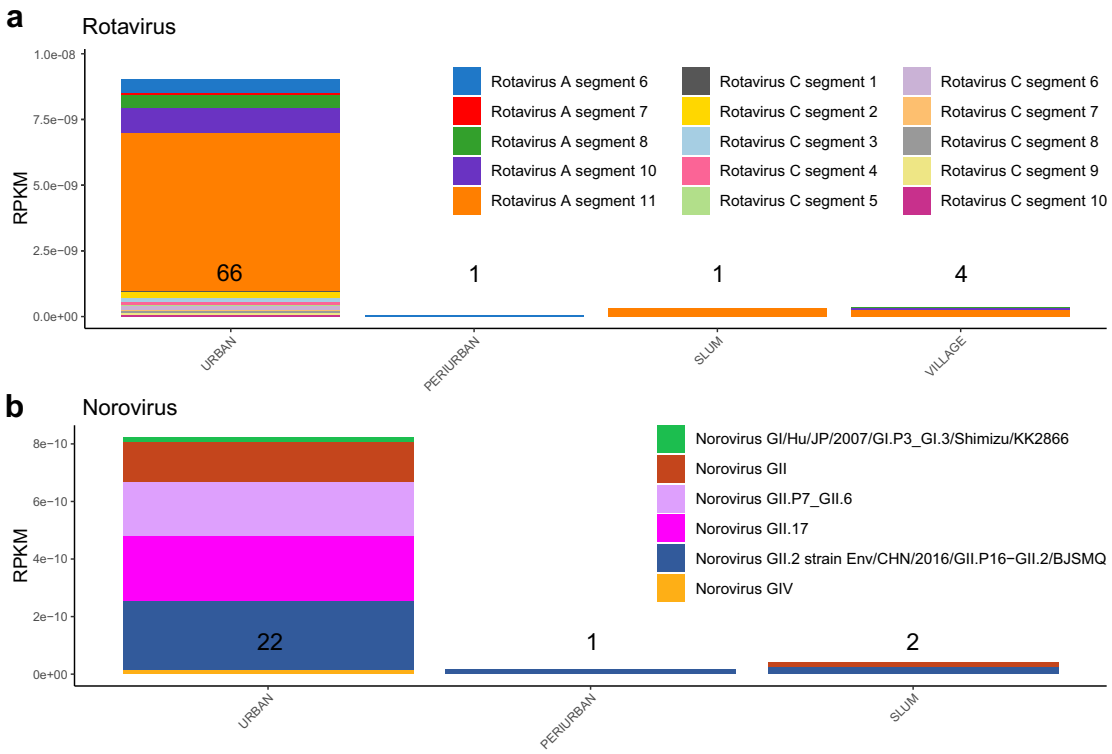

Fig. 4: **Abundance and distribution of the acute gastroenteritis-causing viruses Rotavirus and Norovirus.** The normalised abundances of putative **(a)** Rotavirus species and its genomic segments and **(b)** Norovirus genogroups in untreated sewage samples. The number of samples with positive RPKM values that were accumulated per location are indicated at the base of each bar. RPKM, reads per kilobase per million reads.

virus were both more frequently detected in urban localities (Fig. 5a and b; Chi-squared tests, $P < 0.0001$ and $P = 0.002$, respectively). However, statistical significance was not achieved when analysing the cumulative normalised abundance of astroviruses and saffold virus by locales (Bonferroni adjusted ANOVA tests, $P = 0.9$ and 0.42, respectively). Husavirus and aichi virus were only identified in urban areas and were infrequently detected compared to other human-infecting viruses (Fig. 5c and e). Interestingly, a single rural sample, CR8, contained the measles virus with a cumulative abundance greater than all aichi viral detections combined (Fig. 5d). The possibility that this observation was an artifact due to our RNA-Seq and bioinformatic pipeline was rejected, as it was determined that 30.5% of the 15.9 kb measles virus genome was covered by sequencing reads.

Untreated wastewater provides a glimpse of endemic zoonotic viruses

Chikungunya virus (CHIKV) was identified in 11 urban, 4 peri-urban, and 5 village-associated untreated wastewater endemic samples (Chi-squared test, $P < 0.0001$), with the normalised abundance of CHIKV significantly greater in Nagpur rural areas (Fig. 6a; Bonferroni adjusted ANOVA test, $P = 0.048$). The levels of CHIKV were substantial in several untreated wastewater samples, with four samples having >0.1% of the total RNA-Seq reads mapping across >90% of CHIKV's 11.8 kb genome.

Considering the emergence and cross-species transmission of animal viruses into human hosts, the detection of Jingmen tick virus (JMTV) in untreated wastewater samples was further investigated (Fig. 6b). The frequency of detection of JMTV was greater in urban locales (Chi-squared test, $P < 0.0001$), although the cumulative normalised abundance was non-significant between the collection locations of the untreated wastewater samples (Bonferroni adjusted ANOVA test, $P = 0.5$). Only genomic segment 2 (of four genomic segments) of JMTV was detected in the RNA-Seq data of samples.

Stray dogs and rabies are a significant problem in India, accounting for 35% of the global disease incidents.[25] Like JMTV, rabies virus was more frequently detected in urban localities (Fig. 6c; Chi-squared test, P-value = 0.003) although its cumulative normalised abundance was not significantly different between urban and rural environments (Bonferroni adjusted ANOVA test, $P = 0.11$). Given the increased public awareness of zoonotic diseases and the potential for emergent viruses through human-animal interactions, we included porcine endogenous retrovirus (PERV) E in

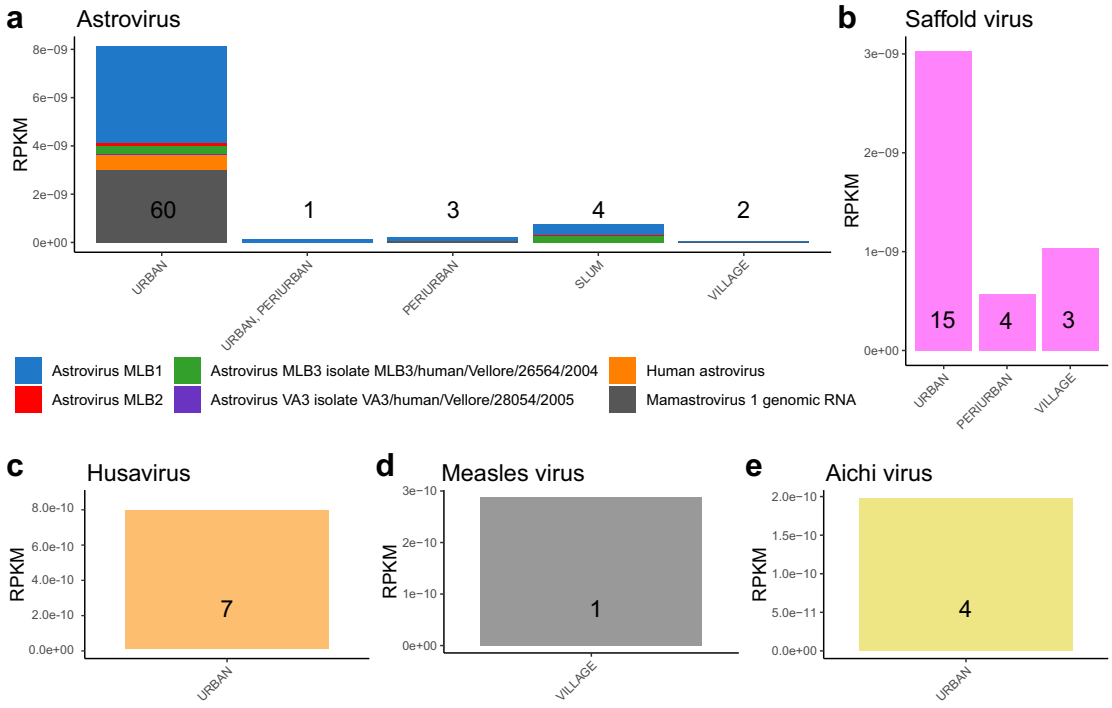

***Fig. 5:* Abundance and distribution of additional noteworthy human pathogenic viruses.** The normalised abundances of putative **(a)** Astrovirus genotypes, **(b)** Saffold virus, **(c)** Husavirus, **(d)** Measles virus, and **(e)** Aichi virus. The number of samples with positive RPKM values that were accumulated per location are indicated at the base of each bar. RPKM, reads per kilobase per million reads.

our analysis (Fig. 6d). No significant difference was observed in the frequency of detection (Fisher's exact test, $P = 1.0$) or cumulative normalised abundance (Bonferroni adjusted ANOVA test, $P = 0.37$) of PERV in the untreated wastewater samples between geographic areas.

Finally, encephalomyocarditis virus (EMCV) was the fourth most frequently detected virus in the RNA-Seq data analysed, abundant in 60 samples (48 urban vs 12 rural; Chi-Squared test, $P < 0.0001$). However, an assessment of its RPKM cumulative abundance shows EMCV as only the 29th most abundant virus detected.

Discussion

This is one of the few studies to utilise culture-and/or probe-independent unbiased RNA sequencing to identify human-infecting RNA viruses in wastewater samples, and to our knowledge, the first which has analysed Indian wastewater samples by this technique. Our findings largely concur with those reported by Corpuz et al. and Gholipour et al. who summarised the main pathogenic viruses detected in wastewater, which has been shown to reflect the pattern of infection in the human population.[26,27] For SARS-CoV-2, we observed a significantly higher normalised abundance in the rural samples compared to the urban samples, coupled with a weak but significant negative association between SARS CoV-2 and humidity. Rural regions in India often have limited healthcare infrastructure, lack of health education, and lower rates of vaccination, which may partially account for this observation, together with pandemic denial and social stigma. Meteorological factors, such as temperature and humidity have previously been suggested to be associated with the transmissibility of certain infectious diseases and pandemic spreading. Here, RNA-Seq data demonstrated a negative correlation between SARS-CoV-2 and humidity, which is concordant with recent cross-sectional surveys in China and the US which showed robust negative correlations between temperature/relative humidity and COVID-19 transmissibility.[28] Although we did not find any significant correlation between SARS-CoV-2 and temperature in the present study, high temperatures and high relative humidity have previously been reported to reduce viability of SARS coronavirus.[29] We are aware of only one earlier clinical study which has examined the incidence of COVID-19 in Nagpur urban region only. During the study period from May to November 2020, the overall COVID-19 RT-PCR positivity rate in the tested samples was 34% in Nagpur region.[30]

Our genomic results demonstrate that HCV was the most frequently detected RNA virus in the untreated wastewater samples and co-occurred with SARS-CoV-2 in 77 samples. Whilst there is no HCV national

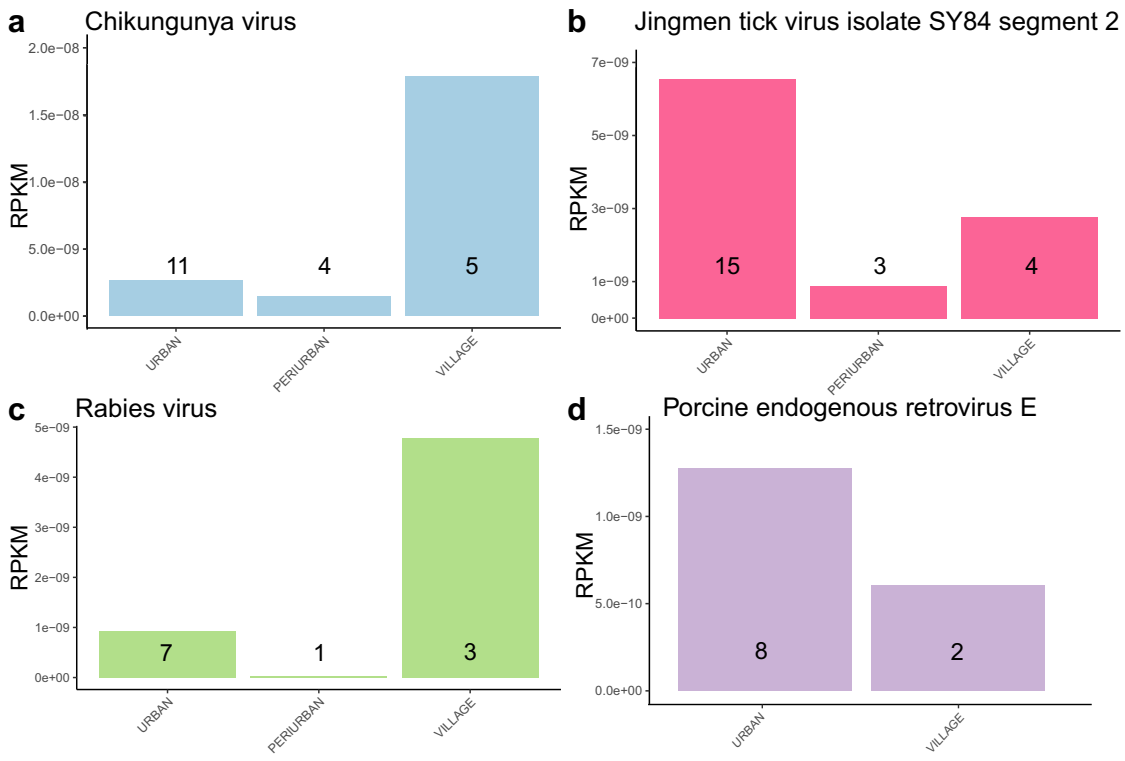

Fig. 6: **Abundance and distribution of human disease-causing zoonotic viruses.** The normalised abundances of putative **(a)** Chikungunya virus, **(b)** Jingmen tick virus, **(c)** Rabies virus, and **(d)** Porcine endogenous retrovirus E. The number of samples with positive RPKM values that were accumulated per location are indicated at the base of each bar. RPKM, reads per kilobase per million reads.

registry in India and thus epidemiology is scarce, HCV seroprevalence has been reported to range between 0.09% and 5.2% across various Indian states.[31–33] The most prominent HCV genotype was genotype 2 which contrasts with findings from a previous epidemiological analysis of HCV which showed that genotype 3 was the most prevalent in India.[34] One possible biological explanation for the concurrent identification of HCV and SARS-CoV-2 is suggested by emerging evidence which shows that HCV increases the expression of the SARS-CoV-2 entry receptor ACE2 in hepatocytes, leading to enhanced entry of both HCV and SARS-CoV-2 into hepatocytes.[35] Furthermore, transmembrane protease serine 2 (TMPRSS2) expression is over-expressed in patients with HCV and TMPRSS2 can potentiate SARS-CoV-2 viral entry.[36] However, the degree and nature of HCV-SARS-CoV-2 interactions and the underlying mechanisms remain largely unknown and require further attention. It remains unexplained why we did not detect Hepatitis A, E, or B in the sequencing data, particularly when such viruses have been reported to be endemic in India.[32,37,38] RNA-Seq analysis also detected EMCV in the wastewater samples. As EMCV is infrequently associated with human illnesses, its detection in untreated wastewater is potentially the result of rodent and/or animal excrement. Nonetheless, EMCV in non-human primates principally results in death through laboured respiration and acute heart failure.[39]

The transcriptomic data also revealed the presence of acute gastroenteritis-causing viruses including several species of type A and B rotaviruses, which were more prevalent in the urban areas, and various norovirus genotypes. These findings are consistent with other next-generation sequencing studies recently summarised in a systematic review which reports the occurrence of several pathogenic viruses in sewage sludge.[27] In India, several enteric viruses have been reported to cause outbreaks, including the co-existence of multiple viruses. Single viral infections showed predominance of enterovirus and rotavirus A (RVA). Genotyping of the viruses revealed predominance of RVA G2P [4], rotavirus group B (RVB) G2 (Indian Bangladeshi lineage), norovirus (NoV) GII.4, adenovirus (AdV)-40, human astrovirus (HAstV)-8 and aichivirus (Aiv) B types.[40] RVA has also been sewage-impacted lakes in highly urbanized regions of western India, where there was a correlation between RVA and coliphages.[41] The burden of viral infectious diseases, such as rotavirus, disproportionally affects lower-income countries. While the development of a

vaccine severely curbed the mortality rate associated with rotavirus, there are still an estimated 128,500 to 215,000 global vaccine-preventable deaths annually due to this virus.[42] Children under 5 years of age are at a greater risk of diarrheal-associated RV mortality, particularly where there is poor sanitation, inadequate water quality or water sources not fit for purpose (drinking, bathing), and unattainable medical care.[43] While rotaviruses infect a wide range of mammals, rotavirus group A (RVA) is more frequently associated with diarrhoeal infections in children and Rotavirus group C (RVC) is more commonly associated with animal contamination.[43–45] Unlike rotavirus, norovirus is problematic in both lower and higher-income countries due to the lack of a licenced vaccine, with most people expected to contract the virus five times in their lifetime.[46] There are an estimated 200,000 global deaths annually attributable to norovirus, the most cases and deaths associated with any bacterial or viral diarrhoeal-causing pathogen,[47] with cases and deaths disproportionally affecting children less than 5 years of age.[48]

Alongside the common enteric viruses, several additional diarrhoeal-causing viruses were identified, including astroviruses and saffold virus which were more frequently detected in urban localities of Nagpur district compared with rural areas. However, husavirus and aichi virus were only detected in urban areas. While astroviruses are well documented as causing sporadic and potentially severe cases of gastroenteritis in children and at-risk groups,[49] saffold virus, husavirus, and aichi virus are emergent human-infecting viruses detected with the advent of next-generation sequencing applications.[50–52]

Notably, Chikungunya, an arthropod-borne disease caused by the Chikungunya virus was detected in 19 wastewater samples and was significantly more prevalent in the rural areas of Nagpur. This is the first report describing the detection of CHIKV in wastewater samples from any country. The CHIKV is currently endemic across India, with regional outbreaks linked to mosquito seasonality and human-mosquito habitat interactions.[53] While the majority of CHIKV infections are acute, with a high fever and arthralgia lasting one to two weeks, the painful arthritic symptoms can last months and even years.[54] Additionally, there is growing evidence for CHIKV-associated neurological manifestations such as encephalopathy in adults and children.[55] A recent study in Pune, India reported a fivefold increase in chikungunya seroprevalence from 2009 to 2019, rising from 8.5% to 53.2%.[56] Therefore, considering the human and economic impacts caused by CHIKV, greater surveillance and accurate assessments of its burden are required to motivate vaccine developments or antiviral discoveries.

Genomic segment 2 of JMTV was detected at a higher frequency in urban locales. To the best of our knowledge, JMTV has not previously been reported in humans or in wastewater samples, underscoring the power of wastewater epidemiology, although JMTV and related viruses have been identified in cattle, rodents, and primates in multiple regions globally.[57] Therefore, further investigations are warranted to identify and perform full genomic sequencing of JMTV or related viruses that are endemic to the district of Nagpur.

Although we detected PERV E-type retrovirus in our analysis, we observed no significant difference in frequency or cumulative normalised abundance between different geographic localities. Currently, specific subtypes of PERV are capable of infecting porcine and human cells,[58] although the principal fear around human PERV infections is currently focused on the risks following xenotransplantation.[59]

EMCV virus has previously been reported in India where it can cause a human febrile illness, although there are no recent epidemiological surveys.[60,61]

Finally, the notable absence of poliovirus may be attributed to the elimination of poliovirus in India in 2011 with the global immunisation campaigns and continuation of vaccination efforts thereafter. Indeed, in India, environmental surveillance complementary to active surveillance of acute flaccid paralysis (AFP) cases played a crucial role in poliovirus elimination and has stimulated COVID-19 surveillance in sewage.[62] Nonetheless, given the long-drawn effort to eliminate polio, wastewater surveillance is likely to play an integral part of public health programs to identify any unreported cases due to lack of notification.

Our study has some limitations and is partially a reflection of the logistics with sampling, processing, and analysing during a pandemic, and the lack of a fully developed infrastructure for WBE. Samples were collected over a relatively short time frame of 8 weeks in one Central Indian district only, and sampling was cross-sectional rather than longitudinal, meaning there was no opportunity to evaluate wastewater variability in terms of composition or to undertake spatiotemporal distribution analyses. Due to resource limitations, personnel shortages during the pandemic, long distances to travel between remote sites, and high temperatures, it was not practical or feasible to deploy autosamplers to undertake 24-h continuous wastewater sampling or to assess other wastewater metrics that are informative of the fate of viruses and other microbes such as flow rate, pH, and turbidity. Also, our sampling of wastewater samples was predominantly from urban areas which might have introduced some bias in the viruses reported. For example, more rural sites might have picked up more agricultural/animal signals. However, the lack of infrastructure in rural settings further complicates the urban/rural comparison. We could not correlate the RNA-Seq results with local clinical incident data since this was not available for enteric and zoonotic viruses. Future studies which aim to build

upon these observations by performing multicentre longitudinal sampling over many months to capture spatiotemporal dynamics. Such studies/monitoring programs would implicitly then capture the impact of seasonality and changes in human behaviour on viral detection. Finally, due to the high levels of variability of metagenomic reads mapping to reference viral sequences, metagenomic assembled genomes from individual samples and genotyping analyses of a conserved phylogenetic marker were not possible.

In conclusion, our analysis provides, for the first time, the detection of several zoonotic RNA viruses in urban and rural wastewater samples in Nagpur India, such as chikungunya and rabies viruses. The COVID-19 pandemic has exposed the weaknesses in rural healthcare in India and other resource-limited settings. This emphasises the need to shift the urban-centric focus of policy makers and researchers to ensure equitable healthcare access and vaccination uptake across all communities, including those most marginalised. Considering the COVID-19 pandemic is an emerging infectious disease of probable animal origin, there is a clear impetus to improve the surveillance of infectious diseases and the sanitary conditions of people worldwide, as well as undertake more intensive surveillance of wildlife species, which pose a danger for emerging zoonotic viruses. Our findings add to the growing body of evidence which suggests that epidemiological surveillance through wastewater-based sequencing can assist in the identification of endemic and problematic viruses at a population scale. Comprehensive application of this effective and adaptable epidemiological alert tool can be leveraged for the rapid assessment of emerging threats, aid pandemic preparedness, and to monitor progress in attaining global health and the Sustainable Development Goals (SDGs) formulated by the United Nations. To realise the full potential reach of this complementary population health tool in India and other LMICs, it will be important to build capacity and scale up current wastewater surveillance and infrastructure beyond SARS-CoV-2 and polio to other pathogens which will also help understand major pathways and drivers of antimicrobial resistance in the environment and zoonotic diseases. However, it will be imperative to standardise and validate existing surveillance methods, including lab procedures, sample size estimations and statistical tests, where normative bodies such as the World Health Organisation and Centers for Disease Control and Prevention alongside national bodies can play a major role in ensuring quality control.

Contributors

TM was the chief investigator. TM, RK and SA conceived and designed the study, had full access to all the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. SS led the RNA-Seq analysis and compiled the results and data interpretation with assistance from AB and UV. AN (Nagpur team member), AH, RN and HD were responsible for collection and transportation of wastewater samples in Nagpur district as well as data compilation. AN (Jaipur team member) contributed to protocol standardization and supervised sample procurement, pre-processing, nucleic acid extraction and shipments to Eurofins. EM, VS and SS assisted with sample cataloguing, storage and pre-processing of wastewater samples. CH advised SS in RNA-Seq data analysis and interpretation. TM, SS and RSK drafted the paper. AS, RG, EA, SC, TB, AT and PM provided advice on wastewater sampling strategy, statistical analysis and general data interpretation. All authors read and edited the manuscript. All authors approved the final version, had full access to all the data, and had final responsibility to submit for publication.

Data sharing statement

Date reported in the manuscript will be made available to investigators who provide a methodologically sound proposal to the corresponding author. The protocol is available upon request. The metadata, processed sequencing data, and scripts required to generate the images and interpret the results of this study are provided as Supplementary Data. The raw RNA-Seq sequencing data analysed in this study is available through NCBI BioProject accession code: PRJNA842541.

Declaration of interests

This study was principally funded by the Global Challenges Research Fund (GCRF) awarded to chief investigator TM and GCRF grant co-investigators RSK, SA, PM, AT, AS, RG, EA, TB, UV, AA, AH as supported by Research England under UKRI. SS and CH received financial support from the Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2273, a Science Foundation Ireland's Spokes Programme which is co-funded under the European Regional Development Fund under Grant Number SFI/14/SP APC/B3032, and a research grant from Janssen Biotech, Inc.

The authors declare no conflict of interest, financial or otherwise.

Acknowledgments

We thank Lorraine Draper for providing intellectual support to SS. We also thank the Nagpur Municipal Cooperation for their assistance with wastewater collection. Finally, we thank Saumitra Mardikar, Bishwadeep, Samiksha, Ankita Bharshankh and Prachi Gandhi for their assistance with wastewater sample collection, transportation of samples and the literature survey.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lansea.2023.100205.

References

1 Mao K, Zhang K, Du W, Ali W, Feng X, Zhang H. The potential of wastewater-based epidemiology as surveillance and early warning of infectious disease outbreaks. *Curr Opin Environ Sci Health*. 2020;17:1–7.
2 Castro-Gutierrez V, Hassard F, Vu M, et al. Monitoring occurrence of SARS-CoV-2 in school populations: a wastewater-based approach. *PLoS One*. 2022;17(6):e0270168.
3 Tlhagale M, Liphadzi S, Bhagwan J, et al. Establishment of local wastewater-based surveillance programmes in response to the spread of infection of COVID-19-case studies from South Africa, the Netherlands, Turkey, and England. *J Water Health*. 2022;20(2):287–299.
4 Hemalatha M, Kiran U, Kuncha SK, et al. Surveillance of SARS-CoV-2 spread using wastewater-based epidemiology: comprehensive study. *Sci Total Environ*. 2021;768:144704.
5 Kopperi H, Tharak A, Hemalatha M, et al. Defining the methodological approach for wastewater-based studies—surveillance of SARS-CoV-2. *Environ Technol Innov*. 2021;23:101696.
6 Tharak A, Kopperi H, Hemalatha M, et al. Longitudinal and long-term wastewater surveillance for COVID-19: infection dynamics and zoning of Urban community. *Int J Environ Res Public Health*. 2022;19(5):2697.
7 Chakraborty P, Pasupuleti M, Jai Shankar MR, et al. First surveillance of SARS-CoV-2 and organic tracers in community wastewater

during post lockdown in Chennai, South India: methods, occurrences, and concurrence. *Sci Total Environ*. 2021;778:146252.
8 Arora S, Nag A, Sethi J, et al. Sewage surveillance for the presence of SARS-CoV-2 genome as a useful wastewater-based epidemiology (WBE) tracking tool in India. *Water Sci Technol*. 2020;82(12):2823–2836.
9 Lamba S, Ganesan S, Daroch N, et al. SARS-CoV-2 infection dynamics and genomic surveillance to detect variants in wastewater–a longitudinal study in Bengaluru, India. *Lancet Reg Health Southeast Asia*. 2023;11:100151.
10 Shah S, Gwee SXW, Ng JQX, Lau N, Koh J, Pang J. Wastewater surveillance to infer COVID-19 transmission: a systematic review. *Sci Total Environ*. 2022;804:150060.
11 Wang Y, Jiang X, Liu L, et al. High-resolution and spatial patterns of virome in wastewater treatment systems. *Environ Sci Technol*. 2018;52:18.
12 Gulino K, Rahman J, Badri M, et al. Initial mapping of the New York City wastewater virome. *mSystems*. 2020;5(3):e00876–e00919.
13 Garcia-Fontana C, Rodriguez-Sanchez A, Munoz-Palazon B, et al. Profile of the spatial distribution of the human and bacteriophage virome in a wastewater treatment plant located in the South of Spain. *Water*. 2020;12(8):2316.
14 Guajardo-Leiva S, Chnaiderman J, Gaggero A, et al. Metagenomic insights into the sewage RNA virosphere of a large city. *Viruses*. 2020;12:1050.
15 Nieuwenhuijse DF, Oude Munnink BB, Phan MVT, et al. Setting a baseline for global urban virome surveillance in sewage. *Sci Rep*. 2020;10:13748.
16 Adriaenssens EM, Farkas K, McDonald JE, Jones DL, Allison HE, McCarthy AJ. Tracing the fate of wastewater viruses reveals catchment-scale virome diversity and connectivity. *Water Res*. 2021;203:117568.
17 Baaijens JA, Zulli A, Ott IM, et al. Variant abundance estimation for SARS-CoV-2 in wastewater using RNA-Seq quantification. *medRxiv*. 2021. https://doi.org/10.1101/2021.08.31.21262938.
18 Balleste E, Blanch AR, Muniesta M, et al. Bacteriophages in sewage: abundance, roles, and applications. *FEMS Microbes*. 2022;3:1–12.
19 Singh KS, Paul D, Gupta A, et al. Indian sewage microbiome has unique community characteristics and potential for population-level disease predictions. *Sci Total Environ*. 2023;858:160178.
20 Heaton KW, Radvan J, Cripps H, et al. Defecation frequency and timing, and stool from in the general population: a prospective study. *Gut*. 1992;33(6):818–824.
21 Specimen Collection. *Packaging and transport guidelines for 2019 nCOV-acute respiratory disease control*. Government of India; 2020. http://ncdc.gov.in/WriteReadData/1892s/50471431021580628750.pdf. Cited November 2022.
22 Arora S, Nag A, Kalra A, et al. Successful application of wastewater-based epidemiology in prediction and monitoring of the second wave of COVID-19 with fragmented sewerage systems—a case study of Jaipur (India). *Environ Monit Assess*. 2022;194(5):342.
23 Okamoto H, Kurai K, Okada S-I, et al. Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes. *Virology*. 1992;188:331–341.
24 Kahn G, Fitzwater S, Tate J, et al. Epidemiology and prospects for prevention of rotavirus disease in India. *Indian Pediatr*. 2012;49:467–474.
25 John D, Royal A, Bharti O. Burden of illness of dog-mediated rabies in India: a systematic review. *Clin Epidemiology Glob Health*. 2021;12:100804.
26 Corpuz MVAC, Buonerba A, Vigliotta G, et al. Viruses in wastewater: occurrence, abundance, and detection methods. *Sci Total Environ*. 2020;745:140910.
27 Gholipour S, Ghalhari MR, Nikaeen M, Rabbani D, Pakzad P, Miranzadeh MB. Occurrence of viruses in sewage sludge: a systematic review. *Sci Total Environ*. 2022;824:153886.
28 Wang J, Tang K, Feng K, et al. Impact of temperature and relative humidity on the transmission of COVID-19: a modelling study in China and the United States. *BMJ Open*. 2021;11:e043863.
29 Chan KH, Peiris JSM, Lam SY, Poon LLM, Yuen KY, Seto WH. The effects of temperature and relative humidity on the viability of the SARS coronavirus. *Adv Virol*. 2011;2011:1–7.
30 Tanna J, Singha B, Nayak AR, et al. Incidence and epidemiological study of COVID-19 in Nagpur Urban region (India) using molecular testing. *medRxiv*. 2021. 2021.5.11.21256719 https://dpi.org/10.1101/2021/05.11.21256719.
31 Puri P, Anand AC, Saraswat VA, et al. Consensus statement of HCV task force of the Indian national association for study of the liver (INASL). Part I: status report of HCV infection in India. *J Clin Exp Hepatol*. 2014;4(2):106–116.
32 Anand A, Shalimar. Hepatitis C virus in India: challenges and successes. *Clin Liver Dis*. 2021;18(3):150–154.
33 Bhaumik P. Epidemiology of viral hepatitis and liver diseases in India. *Euroasian J Hepatogastroenterol*. 2015;5(1):34–36.
34 Shah SR, Rao PN, Sarin SK, et al. Chronic hepatitis C virus infection in India: regional demographics and distribution of viral genotypes. *Indian J Gastroenterol*. 2016;35:469–477.
35 Domovitz T, Ayoub S, Werbner M, et al. HCV infection increases the expression of ACE2 receptor, leading to enhanced entry of both HCV and SARS-CoV-2 into hepatocytes and a coinfection state. *Microbiol Spectr*. 2022;10(6):e0115022.
36 Ronderos D, Omar AMS, Abbas H, et al. Chronic hepatitis-C infection in COVID-19 patients is associated with in-hospital mortality. *World J Clin Cases*. 2021;9(29):8749–8762.
37 Bansal Y, Singla N, Garg K, et al. Seroprevalence of hepatitis A and hepatitis E in patients at a teaching hospital of Northern India over a period of 8 years. *J Family Med Prim Care*. 2022;11(2):567–572.
38 Reddy DCA. Elimination of viral hepatitis: evolution and India's response. *Indian J Public Health*. 2019;63(4):275–276.
39 Carocci M, Bakkali-Kassimi L. The encephalomyocarditis virus. *Virulence*. 2012;3:351–367.
40 Chitambar S, Gopalkrishna V, Chhabra P, et al. Diversity in the enteric viruses detected in outbreaks of gastroenteritis from Mumbai, Western India. *Int J Environ Res Public Health*. 2012;9(3):895–915.
41 Pisharaody L, Suresh S, Mukherji S. Surveillance and seasonal correlation of rotavirus A with coliphages and coliforms in two sewage impacted lakes in highly urbanized regions of Western India. *Environ Sci: Water Res Technol*. 2022;8:139–150.
42 Cates JE, Tate JE, Parashar U. Rotavirus vaccines: progress and new developments. *Expert Opin Biol Ther*. 2022;22:423–432.
43 Varghese T, Kang G, Steele AD. Understanding rotavirus vaccine efficacy and effectiveness in countries with high child mortality. *Vaccines*. 2022;10:346.
44 Rahman M, Banik S, Faruque ASG, et al. Detection and characterization of human group C rotaviruses in Bangladesh. *J Clin Microbiol*. 2005;43:4460–4465.
45 Chepngeno J, Diaz A, Paim FC, Saif LJ, Vlasova AN. Rotavirus C: prevalence in suckling piglets and development of virus-like particles to assess the influence of maternal immunity on the disease development. *Vet Res*. 2019;50:84.
46 Mans J. Norovirus infections and disease in lower-middle and low-income countries, 1997–2018. *Viruses*. 2019;11:341.
47 Pires SM, Fischer-Walker CL, Lanata CF, et al. Aetiology-specific estimates of the global and regional incidence and mortality of diarrhoeal diseases commonly transmitted through food. *PLoS One*. 2015;10:e0142927.
48 Bartsch SM, Lopman BA, Ozawa S, Hall AJ, Lee BY. Global economic burden of norovirus gastroenteritis. *PLoS One*. 2016;11:e0151219.
49 Oude Munnink B, Van der Hoek L. Viruses causing gastroenteritis: the known, the new and those beyond. *Viruses*. 2016;8:42.
50 Yamashita T, Sugiyama M, Tsuzuki H, Sakae K, Suzuki Y, Miyazaki Y. Application of a reverse transcription-PCR for identification and differentiation of aichi virus, a new member of the picornavirus family associated with gastroenteritis in humans. *J Clin Microbiol*. 2000;38:2955–2961.
51 Jones MS, Lukashov VV, Ganac RD, Schnurr DP. Discovery of a novel human picornavirus in a stool sample from a pediatric patient presenting with fever of unknown origin. *J Clin Microbiol*. 2007;45:2144–2150.
52 Oude Munnink BB, Cotten M, Deijs M, et al. A novel genus in the order Picornavirales detected in human stool. *J Gen Virol*. 2015;96:3440–3443.
53 The Translational Research Consortia (TRC) for Chikungunya Virus in India. Current status of chikungunya in India. *Front Microbiol*. 2021;12:695173.
54 Silva LA, Dermody TS. Chikungunya virus: epidemiology, replication, disease mechanisms, and prospective intervention strategies. *J Clin Invest*. 2017;127:737–749.
55 Brizzi K. Neurologic manifestation of chikungunya virus. *Curr Infect Dis Rep*. 2017;19:6.

56 Tomar SJ, Alagarasu K, More A, et al. Decadel change in seroprevalence of chikungunya virus infection in Pune city, India. *Viruses*. 2022;14(5):998.

57 Guo J-J, Lin X-D, Chen Y-M, et al. Diversity and circulation of Jingmen tick virus in ticks and mammals. *Virus Evol*. 2020;6:veaa051.

58 Denner J. Recombinant porcine endogenous retroviruses (PERV-A/C): a new risk for xenotransplantation? *Arch Virol*. 2008;153:1421–1426.

59 Łopata K, Wojdas E, Nowak R, Łopata P, Mazurek U. Porcine endogenous retrovirus (PERV)—molecular structure and replication strategy in the context of retroviral infection risk of human cells. *Front Microbiol*. 2018;9:730.

60 Paul SD, Pavri KM, D'Lima LV, Rajagopalan PK, Goverdhan MK, Singh KR. Isolation of encephalomyocarditis (EMC) virus strains in India. *Indian J Med Res*. 1968;56(3):264–274.

61 Oberste MS, Gotuzzo E, Blair P, et al. Human febrile illness caused by encephalomyocarditis virus infection, Peru. *Emerg Infect Dis*. 2009;15(4):640–646.

62 Ishtiaq F. India must scale up wastewater analysis for health. *Nature India*. 2022. https://doi.org/10.1038/d44151-022-00130-5.