**Evaluation of the impact of concentration and extraction

2 methods on the targeted sequencing of human viruses from**
 Wastewater

Minxi Jiang^a, Audrey L.W. Wang^a, Nicholas A. Be^b, Nisha Mulakken^c, Kara L. Nels

3 **wastewater**
4 Minxi Jiang^a, Audrey
5 a. Department of C
6 b. Physical and Life
7 c. Computing and CA, USA

-
- **methods on the targeted sequencing of human viruses from

Wastewater**

Minxi Jiang^a, Audrey L.W. Wang^a, Nicholas A. Be^b, Nisha Mulakken^c, Kara L. Nelson^a, Rose S. Kantor^{a*}

a. Department of Civil and Environme
- Minxi Jiang", Audrey L.W. Wang", Nicholas A. Be", Nisha Mulakken", Kara L. Nelson"
a. Department of Civil and Environmental Engineering, University of California, Be
b. Physical and Life Sciences Directorate, Lawrence Live Minxi Jiang", Audrey L.W. Wang", Nicholas A. Be", Nisha Mulakken", Kara L. Nelson", Rose S. Kantor"*

a. Department of Civil and Environmental Engineering, University of California, Berkeley, CA, USA

b. Physical and Life 5 b. Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA,

5 c. Computing and Global Security Directorates, Lawrence Livermore National Laboratory, Livermo

CA, USA

*Corresponding 6 b. Physical and Global Security Directorates, Lawrence Livermore National Laboratory, Livermore,

CA, USA.
 4 b. Physical and Coloration: Rose S. Kantor <u>rkantor@berkeley.edu</u>
 4 b. Physical Abstract COMPOSE Abstract
-
-

7 CA, USA
1 *Corresponding author: Rose S. Kantor <u>rkantor@berkeley.edu</u>
2 **Abstract**
1 Sequencing human viruses in wastewater is challenging due to their low abundance compared to t erry server
8 Extends
8 Correspond
8 Sequencing
8 Cotal microb 9 9 * Corresponding author: Rose 5: Kantor <u>Ramon@berkeley.cda</u>
9 * **Abstract**
9 * Sequencing human viruses in wastewater is challenging due
9 * total microbial background. This study compared the impac
9 * methods (Innova 10 **Abstract**

11 Sequencing I

12 total microb

13 methods (Ini

14 viruses follo

15 profiles Inno

16 richness, res 12 total microbial background. This study compared the impact of four virus concentration/extraction
13 methods (Innovaprep, Nanotrap, Promega, Solids extraction) on probe-capture enrichment for human
14 viruses followed b 13 methods (Innovaprep, Nanotrap, Promega, Solids extraction) on probe-capture enrichment for human
14 viruses followed by sequencing. Different concentration/extraction methods yielded distinct virus
15 profiles. Innovapr 14 viruses followed by sequencing. Different concentration/extraction methods yielded distinct virus
15 profiles. Innovaprep ultrafiltration (following solids removal) had the highest sequencing sensitivity and
16 richness profiles. Innovaprep ultrafiltration (following solids removal) had the highest sequencing sensitivity and
16 richness, resulting in the successful assembly of most near-complete human virus genomes. However, it
17 was les 16 richness, resulting in the successful assembly of most near-complete human virus genomes. However, it
17 was less sensitive in detecting SARS-CoV-2 by dPCR compared to Promega and Nanotrap. Across all
18 preparation met was less sensitive in detecting SARS-CoV-2 by dPCR compared to Promega and Nanotrap. Across all
18 reparation methods, astroviruses and polyomaviruses were the most highly abundant human viruses,
19 and SARS-CoV-2 was rare preparation methods, astroviruses and polyomaviruses were the most highly abundant human viruses,

and SARS-CoV-2 was rare. These findings suggest that sequencing success can be increased by using

20 methods that reduce n 19 and SARS-CoV-2 was rare. These findings suggest that sequencing success can be increased by using

20 methods that reduce non-target nucleic acids in the extract, though the absolute concentration of total

21 extracted 19 and Sarks-Cover-2 means and Sarks-Cover-2 means of the extract, though the absolute concentration of total
19 extracted nucleic acid, as indicated by Qubit, and targeted viruses, as indicated by dPCR, may not be
19 dire extracted nucleic acid, as indicated by Qubit, and targeted viruses, as indicated by dPCR, may not be
directly related to targeted sequencing performance. Further, using broadly targeted sequencing panels
may capture viral 22 directly related to targeted sequencing performance. Further, using broadly targeted sequencing panels
23 may capture viral diversity but risks losing signals for specific low-abundance viruses. Overall, this study
24 h may capture viral diversity but risks losing signals for specific low-abundance viruses. Overall, this study
highlights the importance of aligning wet lab and bioinformatic methods with specific goals when
employing probe-24 highlights the importance of aligning wet lab and bioinformatic methods with specific goals when
25 employing probe-capture enrichment for human virus sequencing from wastewater.
26 **Keywords**
27 Targeted sequencing, pr 25 employing probe-capture enrichment for human virus sequencing from wastewater.
26 **Keywords**
27 Targeted sequencing, probe-capture enrichment, human virus, wastewater-based surveillance,
28 wastewater-based epidemiology

- 26 Keywords
27 Targeted sequencing, probe-capture enrichment, human virus, wastewater-based su
28 wastewater-based epidemiology, virus concentration, nucleic acid extraction
29 Sunonsis (c. 30 words)
-

26 **Keywords**
27 Targeted seque
28 wastewater-ba
29 **Synopsis**
30 Four concentra

28 wastewater-based epidemiology, virus concentration, nucleic acid extraction
29 **Synopsis** (\sim 30 words)
20 Four concentration/extraction methods combined with probe-capture sequencing of human virus
21 Target surve t wastewater-based epidemiology, virus concentration, nucleic acid extraction
 Synopsis (~ 30 WOrds)

Four concentration/extraction methods combined with probe-capture sequencing of human viruses in

raw wastewater were co **Synopsis** (~ 30 words)

29 Four concentration/extraction m

21 raw wastewater were compar

22 performance for human virus det

23 1. Introduction 31 raw wastewater were compared. Innovaprep ultrafiltration with solids removal had the best
performance for human virus detection sensitivity, richness, and recovery of near-complete genomes.
33 1. Introduction

32 performance for human virus detection sensitivity, richness, and recovery of near-complete genomes.
33 1. Introduction
1. Introduction

33 1. Introduction
NOTE: This preprint repo

Wastewater-based epidemiology (WBE), previously employed for monitoring enteric viruses like polio¹,
has been widely applied during the COVID-19 pandemic. In 2020, the US Centers for Disease Control and
Prevention (CDC) 36 Prevention (CDC) launched the National Wastewater Surveillance System (NWSS), to build and
37 coordinate the capacity for WBE as a component of the nationwide monitoring of SARS-CoV-2².
38 Subsequently, groups around 37 coordinate the capacity for WBE as a component of the nationwide monitoring of SARS-CoV-2².
38 Subsequently, groups around the world have expanded WBE to include PCR-based monitoring of known
39 seasonal respiratory coordinate the capacity for WBE as a component of the nationwide monitoring of SARS-CoV-2 ².

Subsequently, groups around the world have expanded WBE to include PCR-based monitoring of known

seasonal respiratory viruse

Subsequential Subsequential Subsequential Contribute to CDC NWSS³.

39 Seasonal respiratory viruses including RSV and Influenza A and B, and new PCR panels are expected to

31 Unlike PCR-based virus quantification, sequ 39 contribute to CDC NWSS³.

39 Unlike PCR-based virus quantification, sequencing of viruses in wastewater has the potential to monitor

32 many human viruses at the genome level simultaneously. Reference-based amplicon 40 contribute to CDC NWSS

41 Unlike PCR-based virus qua

42 many human viruses at the

43 tiled panels such as ARTIC

44 targeted amplicons like the

45 tracking of circulating varia

46 clinical data ^{4, 5} However a many human viruses at the genome level simultaneously. Reference-based amplicon sequencing using

43 tiled panels such as ARTIC SARS-CoV-2⁴, ARTIC HAdV-F41⁵, Swift NormalaseTM Amplicon Panel⁶, or

44 targeted ampl 43 tiled panels such as ARTIC SARS-CoV-2⁴, ARTIC HAdV-F41⁵, Swift NormalaseTM Amplicon Panel⁶, or
44 targeted amplicons like those for the VP1 or VP4 regions of enterovirus^{7, 8}, have enabled subtyping and
45 tra tiled panels such as ARTIC SARS-CoV-2 \cdot , ARTIC HAdV-F41 \cdot , Swift Normalase™ Amplicon Panel \cdot , or

targeted amplicons like those for the VP1 or VP4 regions of enterovirus ^{7, 8}, have enabled subtyping and

track targeted amplicons like those for the VP1 or VP4 regions of enterovirus ^{7, 9}, have enabled subtyping and
tracking of circulating variants and strains, providing evidence that wastewater data aligns with available
clinic 1444 clinical data ^{4, 5}. However, amplicon-based sequencing is limited in its ability to detect novel viruses, due

147 to the challenges of degenerate primer design and multiplexing. In contrast, deep untargeted

148 s clinical data ^{4, 5}. However, amplicon-based sequencing is limited in its ability to detect novel viruses, due
to the challenges of degenerate primer design and multiplexing. In contrast, deep untargeted
sequencing offer 48 sequencing offers a comprehensive view of viral diversity in wastewater $9-11$, but human viruses
49 constitute a minimal fraction of the microbial nucleic acids present in wastewater, approximately
50 0.011% of unique sequencing offers a comprehensive view of viral diversity in wastewater γ ¹², but human viruses
constitute a minimal fraction of the microbial nucleic acids present in wastewater, approximately
50 0.011% of unique rea 49 0.011% of unique reads ¹⁰ or 0.1% of the assembled contigs ¹¹. To increase sequencing coverage of

51 human viruses and to allow the detection of divergent or novel viruses in wastewater, probe-capture

52 enrichme 0.011% of unique reads ¹⁰ or 0.1% of the assembled contigs ¹¹. To increase sequencing coverage of

51 human viruses and to allow the detection of divergent or novel viruses in wastewater, probe-capture
 $\frac{1}{2}$ enri 52 enrichment panels have been adopted from clinical research 12 . Here, probes hybridize to DNA targets in
53 a sample, allowing downstream separation of targets from background DNA. Because probe
54 hybridization all enrichment panels have been adopted from clinical research 4° . Here, probes hybridize to DNA targets in

a sample, allowing downstream separation of targets from background DNA. Because probe

by hybridization allows by bridization allows more mismatches than primer binding during PCR, more divergent sequences may
be enriched by probe capture, potentially including novel relatives of known viruses. Recent studies that
have applied viru Example 1.1 allows more capture, potentially including novel relatives of known viruses. Recent studies that

have applied virus probe-capture panels to wastewater-derived samples reported an increase in the

proportion o 56 have applied virus probe-capture panels to wastewater-derived samples reported an increase in the
57 proportion of viral reads up to 81% compared to untargeted sequencing ¹³. Although probe-capture-
58 bacteriophages proportion of viral reads up to 81% compared to untargeted sequencing 13 . Although probe-capture-
based sequencing enriched human viruses, most of the recovered viral content (> 80%) still consisted of
bacteriophages proportion of viral reads up to 81% compared to untargeted sequencing ²⁵. Although probe-capture-
based sequencing enriched human viruses, most of the recovered viral content (> 80%) still consisted of
bacteriophages an 59 bacteriophages and plant viruses $^{14, 15}$. These findings indicate that probe capture panels are still limited
50 in their ability to enrich target sequences in samples with large amounts of background/non-target
51 bacteriophages and plant viruses 14, 14, 14, 14, 15 These findings indicate that probe capture panels are still limited

in their ability to enrich target sequences in samples with large amounts of background/non-target

s

For the main target sequences in tampies with large amount of background, the distribution
62 in the main viruses.
63 Frior to the COVID-19 pandemic, sequencing-based wastewater virus studies relied on large-volume
64 time 62 of human viruses.

63 Prior to the COVID-19 pandemic, sequencing-based wastewater virus studies relied on large-volume

64 time-intensive methods that had initially been developed to culture infectious viruses (e.g., p Framman viruses.

63 Prior to the COVI

64 time-intensive me

65 glycol precipitatio

66 studies reported is

67 untargeted sequer

68 pandemic the deta time-intensive methods that had initially been developed to culture infectious viruses (e.g., polyethylene

glycol precipitation, skim milk flocculation, ultracentrifugation, and membrane filtration). Multiple

studies re glycol precipitation, skim milk flocculation, ultracentrifugation, and membrane filtration). Multiple
studies reported that the choice of concentration method influenced the resulting virus profiles by
untargeted sequenci studies reported that the choice of concentration method influenced the resulting virus profiles by

untargeted sequencing $9,16$ and few studies reported any sequences from enveloped viruses. During the

pandemic, the de 67 untargeted sequencing ^{9, 16} and few studies reported any sequences from enveloped viruses. During the
68 pandemic, the demand for rapid routine monitoring of SARS-CoV-2 led to the development and wider
69 adoption of untargeted sequencing ^{9, 16} and few studies reported any sequences from enveloped viruses. During the

pandemic, the demand for rapid routine monitoring of SARS-CoV-2 led to the development and wider

adoption of stream adoption of streamlined concentration/extraction methods with lower sampling volumes, ending with

qPCR or digital PCR quantification^{17, 18}. These methods included size separation (e.g., Innovaprep

ultrafiltration pipe q PCR or digital PCR quantification^{17, 18}. These methods included size separation (e.g., Innovaprep

11 ultrafiltration pipette, centrifugal ultrafiltration), capture based on virus surface characteristics (e.g.,

12 Na qPCR or digital PCR quantification^{27, 26}. These methods included size separation (e.g., Innovaprep

dutrafiltration pipette, centrifugal ultrafiltration), capture based on virus surface characteristics (e.g.,

Nanotrap Nanotrap beads, electron-negative HA membrane), and direct nucleic acid extraction (e.g., Promega
Wizard Enviro large-volume extraction, or extraction of wastewater solids after centrifugation). These
routine monitoring m Wizard Environ-negative HA membrane), and variable matrix members of the tentrifugation). These
routine monitoring methods were also used to obtain SARS-CoV-2 RNA for sequencing, with varying
success $^{19-22}$ and later e The multime monitoring methods were also used to obtain SARS-CoV-2 RNA for sequencing, with varying
success $^{19-22}$ and later extended for detection of a wider spectrum of viruses 14 , 15 , $^{23-25}$. To date, few 75 success $^{19-22}$ and later extended for detection of a wider spectrum of viruses $^{14, 15, 23-25}$. To date, few
76 studies have directly compared the effects of different methods on the success of virus probe-capture studies have directly compared the effects of different methods on the success of virus probe-capture 76 studies have directly compared the effects of different methods on the success of virus probe-capture

(300 µL for direct extraction and 50 mL for HA filtration) and suggested that direct extraction may yield a

19 lower equivalent volume of viruses in the final extracted nucleic acid compared to pre-filtered samples

23. 79 Iower equivalent volume of viruses in the final extracted nucleic acid compared to pre-filtered samples

23. Spurbeck et al (2023) indirectly compared five wastewater virus concentration/extraction methods,

31 but ea ²³. Spurbeck et al (2023) indirectly compared five wastewater virus concentration/extraction methods,

81 but each was applied to wastewater samples from a different location(s). They found that Innovaprep

82 ultrafilt \overline{a} 81 but each was applied to wastewater samples from a different location(s). They found that Innovaprep
82 ultrafiltration yielded the highest virus sequence recovery in untargeted RNA sequencing, although most
83 sequence B2 ultrafiltration yielded the highest virus sequence recovery in untargeted RNA sequencing, although most

sequences corresponded to bacteriophage 24 . These findings highlight the potential impact of

concentration/e sequences corresponded to bacteriophage 24 . These findings highlight the potential impact of
concentration/extraction methods on targeted sequencing of diverse viruses, but direct comparisons
and analysis of potential b sequences corresponded to bacteriophage ²⁴. These findings highlight the potential impact of

84 concentration/extraction methods on targeted sequencing of diverse viruses, but direct comparisons

85 and analysis of pote

Between the magnetic concentration methods on sequencing performance are needed,

84 especially for targeted sequencing using probe-capture panels.

87 In this study, four wastewater virus concentration/extraction methods BREET AND ANY OF PERSONAL BIASE FROM CONCRETENT CONSUMING THE METHOD AS A SERVICE AND A SERVICE AND A SERVICE A

87 and this study, four wastewater virus concentration/extraction methods were selected based on their

88 on Experimited angles and probe-capture panels probe-
88 ongoing use in wastewater surveillance efforts, and the
89 sequencing was compared for each method. The wastewater is
90 resulting nucleic acids were enriched using the ongoing use in wastewater surveillance efforts, and the success of probe-capture enrichment
sequencing was compared for each method. The wastewater input volume was held constant, and the
resulting nucleic acids were enric Bequencing was compared for each method. The wastewater input volume was held constant, and the

resulting nucleic acids were enriched using the Illumina virus surveillance panel (VSP). The evaluation of

methods performan Besulting nucleic acids were enriched using the Illumina virus surveillance panel (VSP). The evaluation of

Separation of methods performance included total nucleic acid quality, unique sequence output, taxonomic

composit methods performance included total nucleic acid quality, unique sequence output, taxonomic

92 composition, richness, recovered genome completeness, and sensitivity comparisons between

94 sequencing and dPCR. Ultimately, 1912 composition, richness, recovered genome completeness, and sensitivity comparisons between
93 sequencing and dPCR. Ultimately, these findings improve our understanding of wet lab approaches and
94 their compatibility w 93 sequencing and dPCR. Ultimately, these findings improve our understanding of wet lab approaches and
94 their compatibility with virus probe-capture enrichment and sequencing, informing tailored responses to
95 emerging their compatibility with virus probe-capture enrichment and sequencing, informing tailored responses to

emerging viral threats.

96 2. Materials and Methods

97 2.1 Sample collection energing viral threats.
95 emerging viral threats.
96 2. Materials and Methods
97 2.1 Sample collection

96 2. Materials and
97 2.1 Sample collection
98 Influent wastewater wa 96 2. Materials and Methods
97 2.1 Sample collection
98 Influent wastewater was collected a
99 and April 26th, 2023, from the EBM
90 serves approximately 700,000 peop 97 2.1 Sample conection
99 and April 26th, 2023, from
99 and April 26th, 2023, from
90 sample was transported
92 coronavirus (BCoV) was and April 26th, 2023, from the EBMUD wastewater treatment plant (Alameda County, CA). This facility
serves approximately 700,000 people, receiving domestic and industrial wastewater. On each date, the
sample was transporte Survey approximately 700,000 people, receiving domestic and industrial wastewater. On each date, the

91 sample was transported to the laboratory on ice, and twelve 40 mL aliquots were prepared. Bovine

92 coronavirus (BC 101 sample was transported to the laboratory on ice, and twelve 40 mL aliquots were prepared. Bovine

102 coronavirus (BCoV) was added to each tube as a sample processing control to assess viral RNA recovery.

103 First, o 102 coronavirus (BCoV) was added to each tube as a sample processing control to assess viral RNA recovery.

103 First, one vial of BCoV (Merck) vaccine powder was resuspended in 2 mL 0.1 mM Tris-

104 Ethylenediaminetetraa 102 coronavirus (BCoV) was added to each tube as a sample processing control to assess viral RNA recovery.

103 First, one vial of BCoV (Merck) vaccine powder was resuspended in 2 mL 0.1 mM Tris-

104 Ethylenediaminetetra

108 pipette concentration (IP method), Nanotrap bead concentration (NT method), Promega large-volume 107 Four concentration and extraction methods (described below) were employed in this study: Innovaprep
108 pipette concentration (IP method), Nanotrap bead concentration (NT method), Promega large-volume
109 direct extrac 2.2 Concentration and extraction
107 Four concentration and extraction methods (described be
108 pipette concentration (IP method), Nanotrap bead conce
109 direct extraction (PMG method), and pelleted solids direct
110 per 2.2 Concentration and extraction
107 Four concentration and extraction me
108 pipette concentration (IP method), N
109 direct extraction (PMG method), and
110 performed on three 40-mL aliquots
111 consisting of 40 mL 1x ph 110 performed on three 40-mL aliquots of wastewater per sample date, alongside a negative control
111 consisting of 40 mL 1x phosphate-buffered saline (PBS) solution (**Table S1**). All methods resulted in 100
112 ul purifie bead concentration (IT method), Manotrap bead concentration (NT method), Fromega large-volume
109 direct extraction (PMG method), and pelleted solids direct extraction (Solids method). Each method was
110 performed on thre and the performed on three 40-mL aliquots of wastewater per sample date, alongside a negative control consisting of 40 mL 1x phosphate-buffered saline (PBS) solution (Table S1). All methods resulted in 100 ul purified tota 111 consisting of 40 mL 1x phosphate-buffered saline (PBS) solution (**Table S1**). All methods resulted in 100 μ purified total nucleic acid (TNA). 111 consisting of 40 mL 1x phosphate-buffered same (PBS) solution (Table S1). All methods resulted in 100

4112 purified total nucleic acid (TNA). 112 µl purified total nucleic acid (TNA).

In the first two methods, separate concentration and extraction steps were applied. In the IP method,

114 400 µL of 5% Tween 20 was added to the wastewater sample and mixed by inversion, followed by

115 concentration at The contribution at 7000 x g for 10 min. The supernatant was ultrafiltered using the automatic HF

concentration Pipette (Innovaprep CP-SelectTM) and eluted with the elution fluid (Innovaprep) to

produce the viral conc Concentration Pipette (Innovaprep CP-SelectTM) and eluted with the elution fluid (Innovaprep) to

117 produce the viral concentrate (ranging from 160 to 882 μ **L, Table S1**). TNA was then extracted from up

118 to 200 117 produce the viral concentrate (ranging from 160 to 882 μL, **Table S1**). TNA was then extracted from up

118 to 200 μL of viral concentrate using the Allprep PowerViral DNA/RNA kit (Qiagen) and eluted in 100 μL,

119 f 117 produce the viral concentrate (ranging from 160 to 882 μL, Table 31). TNA was then extracted from up
118 to 200 μL of viral concentrate using the Allprep PowerViral DNA/RNA kit (Qiagen) and eluted in 100 μL,
119 Micro 118 to 200 μL of viral concentrate using the Allprep PowerViral DNA/RNA Kit (Qiagen) and eluted in 100 μL,
119 following the manufacturer's liquid sample extraction protocol. The NT method followed the Nanotrap®
120 Micro 119 following the manufacturer's liquid sample extraction protocol. The NT method followed the Nanotrap

120 Microbiome A Protocol, compatible with AllPrep PowerViral DNA/RNA Kit (APP-091 December 2022).

121 Briefly, 115 First Protocol, 115 μL of Nanotrap® Enhancement Reagent 2 (ER2) and 600 μL of Nanotrap Microbiome A

122 Particles (Ceres Nanosciences) were sequentially added to each sample, followed by mixing and

123 incubation. The b Farticles (Ceres Nanosciences) were sequentially added to each sample, followed by mixing and

123 incubation. The beads were separated from the solution on a magnetic rack and resuspended in 1 mL of

124 molecular-grade incubation. The beads were separated from the solution on a magnetic rack and resuspended in 1 mL of
124 molecular-grade water, followed by another separation using the magnetic rack. The beads were then
125 mixed with 600 molecular-grade water, followed by another separation using the magnetic rack. The beads were then

125 mixed with 600 µL of preheated PM1 + Beta-mercaptoethanol solution from the Allprep PowerViral kit,

126 and the mixt molecular-grade water, followed by another separation using the magnetic rack. The beads were then
125 mixed with 600 µL of preheated PM1 + Beta-mercaptoethanol solution from the Allprep PowerViral kit,
126 and the mixture 126 mixture was heated at 95°C for 10 min to release nucleic acids. Beads were removed using the
127 magnetic rack and the supernatant was used for subsequent extraction steps using the Allprep
128 PowerViral DNA/RNA kit

magnetic rack and the supernatant was used for subsequent extraction steps using the Allprep
128 PowerViral DNA/RNA kit liquid protocol, resulting in 100 µL of final TNA.
130 The other two methods were direct extractions f 128 PowerViral DNA/RNA kit liquid protocol, resulting in 100 µL of final TNA.

129 The other two methods were direct extractions from either total wastewater or pelleted solids. The
 PMG method used the commercial kit f The other two methods were direct extractions from either total was
 129 The other two methods were direct extractions from either total was
 130 PMG method used the commercial kit from Promega (Wizard® Enviro T

131 m **PMG method** used the commercial kit from Promega (Wizard® Enviro Total Nucleic Acid) following the

manufacturer's protocol. Briefly, 0.5 mL of protease was added to each 40 mL wastewater sample and

incubated for 30 min 131 manufacturer's protocol. Briefly, 0.5 mL of protease was added to each 40 mL wastewater sample and

132 incubated for 30 minutes. After centrifugation at 3000 x g for 10 min, binding buffers and isopropanol

133 were 132 incubated for 30 minutes. After centrifugation at 3000 x g for 10 min, binding buffers and isopropanol

133 were added to the resulting supernatant before passing it through the PureYield^r^m binding column. The

13 The method of the resulting supernatant before passing it through the PureYield™ binding column. The bound nucleic acids were washed and then eluted in 1 mL of nuclease-free water. The eluted samples were further purifie Figure and the resultion of the resultion of the resultion of the resultion of the resulting supernated to the PureYield™ Minicolumn, resulting in a final
135 were further purified, concentrated, and eluted using the Pur were further purified, concentrated, and eluted using the PureYieldTM Minicolumn, resulting in a final

136 total nucleic acid volume of 100 µL. In the **Solids method**, the 40 mL wastewater sample was centrifuged

137 a 136 total nucleic acid volume of 100 µL. In the **Solids method**, the 40 mL wastewater sample was centrifuged
137 at 20,000 x g for 10 minutes to pellet the solids. Total nucleic acid was then extracted from 0.2 g (wet
138 137 at 20,000 x g for 10 minutes to pellet the solids. Total nucleic acid was then extracted from 0.2 g (wet
138 weight) of solid pellets using the Allprep PowerViral DNA/RNA extraction kit. This followed the
139 manufact weight) of solid pellets using the Allprep PowerViral DNA/RNA extraction kit. This followed the

139 manufacturer's solids extraction protocol, which included a 10-minute bead-beating step after the

140 addition of PM1 an manufacturer's solids extraction protocol, which included a 10-minute bead-beating step after the
140 addition of PM1 and Beta-mercaptoethanol solution. The final extracted TNA were eluted in 100 µL of
141 DNA and RNA conc

140 addition of PM1 and Beta-mercaptoethanol solution. The final extracted TNA were eluted in 100 µL of
141 nuclease-free water.
142 DNA and RNA concentrations were quantified using the Qubit 1X dsDNA HS Assay (Fisher Scie nuclease-free water.

142 DNA and RNA concentrations were quantified using the Qubit 1X dsDNA HS Assay (Fisher Scientific) and

143 Qubit RNA HS Assay (Fisher Scientific), respectively. Aliquots of all extracts were stored 142 DNA and RNA concer
143 Qubit RNA HS Assay
144 quantified by dPCR w
145 2.3 Digital PCR qua DNA and RNA concentrations were quantified using the Qubit 1X dsDNA HS Assay (Fisher Scientific) and
143 Qubit RNA HS Assay (Fisher Scientific), respectively. Aliquots of all extracts were stored at -20°C and
144 quantifie

144 quantified by dPCR within one week and at -80℃ for subsequent sequencing library preparation.

145 2.3 Digital PCR quantification of SARS-CoV-2 in the extracted total nucleic acid

146 Digital PCR was performed on the 2.3 Digital PCR quantification of SARS-CoV-2 in the extracted total nucleic acid
146 Digital PCR was performed on the QIAcuity Four Platform Digital PCR System (Qiagen). The de
147 SARS-CoV-2 and BCoV assays' primers, prob 2.3 Digital PCR was performed on the QIAcuity Four Platform Digital PCR System (Qiager
145 Digital PCR was performed on the QIAcuity Four Platform Digital PCR System (Qiager
148 S2a. The reaction mixtures were prepared usi 147 SARS-CoV-2 and BCoV assays' primers, probes, and thermal cycling conditions are summarized in **Table**

148 S2a. The reaction mixtures were prepared using the QIAcuity OneStep Advanced Probe Kit (Qiagen) and

149 loaded 147 SARS-COV-2 and BCOV assays primers, probes, and thermal cycling conditions are summarized in Table

148 S2a. The reaction mixtures were prepared using the QIAcuity OneStep Advanced Probe Kit (Qiagen) and

149 loaded on 32a. The reaction mixtures were prepared using the QIACuity OneStep Advanced 110be Kit (Qiagen) and
149 loaded onto either 8.5k 24-well or 26k 24-well nanoplates (Qiagen). Details of reaction mixture
150 composition and vo 150 composition and volumes were summarized in **Table S2b**. The positive control was linearized gene
151 plasmids from Integrated DNA Technologies, and the negative control was nuclease-free water. See
152 **Figure S1** for 150 composition and volumes were summarized in Table 32b. The positive control was intearized gene
151 plasmids from Integrated DNA Technologies, and the negative control was nuclease-free water. See
152 **Figure S1** for ex 152 Figure S1 for examples of the partition fluorescence plots of positive and negative controls. The number
153 of valid partitions ranged from 7,920 to 8,269 per well for 8.5k plates and 12,548 to 25,493 per well for
153 152 Figure S1 for examples of the partition fluorescence plots of positive and negative controls. The number of valid partitions ranged from 7,920 to 8,269 per well for 8.5k plates and 12,548 to 25,493 per well for 153 of value partitions ranged from 7,920 to 8,269 per well for 8.5k plates and 12,548 to 25,493 per well for
153 to 8,269 per well for 8.5k plates and 12,548 to 25,493 per well for 8.5k per well for 8.5k per well for 8.
1

154 26k plates. Data were analyzed using the QIAcuity Suite Software V1.1.3 (Qiagen, Germany) with
155 automated settings for threshold and baseline, followed by manual inspection. Results were plotted
156 using a customi using a customized Python script. dMIQE checklists ²⁶ are provided in **Table S3**. The operational limit of
157 detection was treated as \geq 3 positive partitions per well.
158 **2.4 Library preparation and targeted seque**

using a customized Python script. dMIQE checklists ²⁶ are provided in Table S3. The operational limit of
157 detection was treated as ≥3 positive partitions per well.
158 **2.4 Library preparation and targeted sequencin** 2.4 Library preparation and targeted sequencing
159 Before library preparation, DNA and RNA quality were n
160 HS NGS Fragment 1-50 kb assay and Bioanalyzer (Agile
161 assay, respectively. Library preparation followed the
 2.4 Library preparation and targeted sequencing
159 Before library preparation, DNA and RNA quality were
160 HS NGS Fragment 1-50 kb assay and Bioanalyzer (Ag
161 assay, respectively. Library preparation followed the
162 m 160 HS NGS Fragment 1-50 kb assay and Bioanalyzer (Agilent 2100) with the Agilent RNA 6000 Pico RNA
161 assay, respectively. Library preparation followed the Illumina RNA Prep with Enrichment kits with
162 modifications t 161 assay, respectively. Library preparation followed the Illumina RNA Prep with Enrichment kits with
162 modifications to total input (Illumina, San Diego, CA, USA). In brief, the mixture of purified DNA and RNA
163 from modifications to total input (Illumina, San Diego, CA, USA). In brief, the mixture of purified DNA and RNA

163 from samples collected on April 19 and April 26 was diluted with nuclease-free water such that the final

164 163 from samples collected on April 19 and April 26 was diluted with nuclease-free water such that the final
164 concentration of RNA was ≤ 100 ng/µl. Dilution was not conducted for IP and NT samples due to the low
16 164 concentration of RNA was ≤ 100 ng/ μ l. Dilution was not conducted for IP and NT samples due to the low
165 RNA concentrations. The DNA and RNA from samples collected on March 1 were used directly as the
166 inpu 165 RNA concentrations. The DNA and RNA from samples collected on March 1 were used directly as the

166 input for library preparation without dilution for all concentration/extraction methods (**Table S1**). Next,

167 8.5-166 input for library preparation without dilution for all concentration/extraction methods (**Table S1**). Next,
167 8.5-µL of each sample was denatured followed by first- and second-strand DNA synthesis. Tagmentation
168 o 167 8.5-µL of each sample was denatured followed by first- and second-strand DNA synthesis. Tagmentation
168 of the total enriched double-stranded cDNA was performed using bead-linked transposons (BLT), and
169 adapter seq 168 of the total enriched double-stranded cDNA was performed using bead-linked transposons (BLT), and
169 adapter sequences were added at the same time. The resulting fragments were purified and amplified
170 to add index 168 of the total enriched double-stranded cDNA was performed using bead-linked transposons (BLT), and
169 adapter sequences were added at the same time. The resulting fragments were purified and amplified
170 to add index 170 to add index sequences. Libraries were quantified using the Qubit dsDNA broad-range Assay Kit.
171 Enrichment was performed with the Illumina VSP Panel by pooling 200 ng of each library from three
172 biological replic 171 Enrichment was performed with the Illumina VSP Panel by pooling 200 ng of each library from three
172 biological replicates into hybridization reactions. This step was followed by bead-based capture of
173 hybridized p biological replicates into hybridization reactions. This step was followed by bead-based capture of

173 by bridized probes, amplification, clean-up, and quantification of the final enriched library. After library

174 pre hybridized probes, amplification, clean-up, and quantification of the final enriched library. After library
174 preparation, all enriched samples were pooled in equimolar ratios and sequenced on one lane of
175 Illumina No 174 preparation, all enriched samples were pooled in equimolar ratios and sequenced on one lane of
175 Illumina Novaseq 6000 SP 150PE.
176 2.5 Bioinformatics Analysis Pipeline
177 Sequence data were quality trimmed using B

175 Illumina Novaseq 6000 SP 150PE.

175 **2.5 Bioinformatics Analysis Pipeline**

177 Sequence data were quality trimmed using BBduk ²⁷ to remove adaptors and filter out low-quality reads

178 and short reads. Seqkit was 2.5 Bioinformatics Analysis Pip
175 Sequence data were quality trimn
178 and short reads. Seqkit was use
179 Before taxonomy classification, 180 176 2.5 Biomformatics Analysis Figure
177 Sequence data were quality trimmed us
178 and short reads. Seqkit was used to d
180 GRCh38.p14 (RefSeq GCF_000001405.4
181 non-human unique reads were class
decontaminated version Sequence data were quality trimmed using BBduk ²⁷ to remove adaptors and filter out low-quality reads

and short reads. Seqkit was used to deduplicate reads and summarize unique reads ²⁸ (Figure 1b).

179 Before taxon and short reads. Seqkit was used to deduplicate reads and summarize unique reads ²⁸ (Figure 1b).

179 Before taxonomy classification, human reads were filtered using bowtie2 (v2.5.1)²⁹ by mapping to

180 GRCh38.p14 (R Before taxonomy classification, human reads were filtered using bowtie2 (v2.5.1) ²⁵ by mapping to
180 GRCh38.p14 (RefSeq GCF_000001405.40) and CHM13v2.0 (RefSeq GCF_009914755.1). The remaining
181 non-human unique reads 181 non-human unique reads were classified by Centrifuge (1.0.4)³⁰ and Recentrifuge³¹ using a decontaminated version of NCBI-nt database (NCBI release date June 5, 2023). A minimum hit length (MHL) threshold of 15 was non-human unique reads were classified by Centrifuge (1.0.4) ³⁰ and Recentrifuge ³¹ using a
182 decontaminated version of NCBI-nt database (NCBI release date June 5, 2023). A minimum hit length
183 (MHL) threshold of 1 183 (MHL) threshold of 15 was employed for Centrifuge. An MHL threshold of 40 was subsequently applied

184 in Recentrifuge for downstream analysis. After classification, all viral reads were extracted from each

185 samp 184 in Recentrifuge for downstream analysis. After classification, all viral reads were extracted from each

185 sample using rextract and viral sequence similarities between samples were compared using MASH ³².

186 Pa 185 sample using rextract and viral sequence similarities between samples were compared using MASH ³².

186 Pairwise Mash distances were calculated for the construction of the PCoA plot using the

187 sklearn.decomposit sample using rextract and viral sequence similarities between samples were compared using MASH 22 .

186 Pairwise Mash distances were calculated for the construction of the PCoA plot using the

187 sklearn.decompositio Sklearn decomposition PCA package in Python. A PERMANOVA test with 999 permutations was
188 performed using the vegan package (2.6.4) in R³³. One sample (PMG_426_2) displayed distinct
189 sequence properties from the ot 188 performed using the vegan package (2.6.4) in R³³. One sample (PMG_426_2) displayed distinct
189 sequence properties from the other two biological replicates in the original PCoA (**Figure S2**) and yielded
190 unexpec performed using the vegan package (2.6.4) in R 33 . One sample (PMG_426_2) displayed distinct
189 sequence properties from the other two biological replicates in the original PCoA (**Figure S2**) and yielded
190 unexpect sequence properties from the other two biological replicates in the original PCOA (Figure 32) and yielded
190 unexpectedly low unique read counts (Table S1), likely due to unsuccessful enrichment during the
191 library pre Interactive of the countries of the S1), they due to different during the library preparation. This sample was excluded from all sequencing analyses. To precisely identify SARS-
CoV-2 reads during the sampling period, uniq 193 severe acute respiratory syndrome-related coronavirus (taxID: 694009) were extracted and mapped to
194 references from the GISAID database 34 downloaded on January 2, 2024 (**Table S4**). The references 194 references from the GISAID database 34 downloaded on January 2, 2024 (Table S4). The references 194 Freferences from the GISAID database ³⁴ downloaded on January 2, 2024 (**Table S4**). The references trom the GISAID database ³⁴ downloaded on January 2, 2024 (**Table S4**). The references

196 January 1, 2023, to May 31, 2023. The mapped reads were subjected to additional filtering using

197 reformat.sh from BBduk ²⁷. Mapped reads with fewer than 5 mismatches were considered as classified

198 To determi

197 Terms of the BBduk²⁷. Mapped reads with fewer than 5 mismatches were considered as classified
198 SARS-CoV-2 reads.
199 To determine putative virus host assignments for each read (Figure 2c and 2d), the NCBI taxonomy reformat.sh from BBduk ²⁷. Mapped reads with fewer than 5 mismatches were considered as classified
198 SARS-CoV-2 reads.
199 To determine putative virus host assignments for each read (Figure 2c and 2d), the NCBI taxonom 199 To determine puta
200 database ³⁵, which
201 by Recentrifuge, ar
202 of DNA and RNA v
203 results. To focus s 199 To determine putative virus host assignments for each read (Figure 2c and 2d), the NCBI taxonomy

200 database ³⁵, which includes virus host information, was queried with the NCBI taxID of the best hit given

201 by database ³⁵, which includes virus host information, was queried with the NCBI taxID of the best hit given
201 by Recentrifuge, and the results were manually inspected (see SI methods for details). The comparison
202 of D 202 of DNA and RNA viruses was conducted at the virus kingdom level based on taxonomic classification
203 of DNA and RNA viruses was conducted at the virus kingdom level based on taxonomic classification
203 occurred at th results. To focus specifically on human viruses, the identification of all human viruses (section 3.3)
204 occurred at the species level based on NCBI taxonomy ³⁵. Species-level richness was determined by
205 counting th occurred at the species level based on NCBI taxonomy ³⁵. Species-level richness was determined by

205 counting the unique human virus species, with a cutoff of > 10 classified reads applied to discard low-

206 abundanc 204 occurred at the species level based on NCBI taxonomy \degree . Species-level richness was determined by

205 counting the unique human virus species, with a cutoff of > 10 classified reads applied to discard low-

206 a

205 counting the unique term of the unique species, where $\frac{37}{2}$ were further subjected to quality filtering requiring a length of > 1000 bp and 207 abundance viruses (Figure 35a and 3b). Some reads were assigned to species-level NCBI taxibs that
207 paradoxically lacked clear species-level taxonomy names in the database. To make this apparent, these
208 species a 208 species are displayed with "unclassified" appended to the species name (see SI methods for details).

209 To maximize the recovery of near-complete viral genomes for wet lab method comparison, all unique

210 reads we 208 species are displayed with "unclassified" appended to the species name (see SI methods for details).

209 To maximize the recovery of near-complete viral genomes for wet lab method comparison, all unique

210 reads we The recoverage the symbology and the recoverage of $>$ 10x. These filtered viral genome complete viral genome complete viral genome complete viral genome comparison and an average coverage of $>$ 10x. These filtered assem reads were separately assembled using SPAdes with the -meta option (v3.15.5) ³⁶. All virus scaffolds

identified by VirSorter2³⁷ were further subjected to quality filtering, requiring a length of > 1000 bp and

an ave identified by VirSorter2 ³⁷ were further subjected to quality filtering, requiring a length of > 1000 bp and

212 an average coverage of > 10x. These filtered assemblies were then subjected to BLASTn search against

213 213 **EXECT ASSEM INTERT COVER THE SETT ASSEM** THE NCBI nt virus database, with stringent quality filters applied: > 80% identity, > 90% alignment/query length, and an e-value < 1E-8. The best hit based on bitscore was ret 214 length, and an e-value < 1E-8. The best hit based on bitscore was retained for each assembled scaffold

215 and information including virus name, taxlD, genome completeness, and genome length was retrieved

216 from N 215 and information including virus name, taxlD, genome completeness, and genome length was retrieved

216 from NCBI via the dataset and dataformat functions ³⁸. Hit genomes were retained only if complete, and

217 assem 216 from NCBI via the dataset and dataformat functions 38 . Hit genomes were retained only if complete, and
217 assembled genomes were used for further analysis if the alignment length was > 70% of the complete
228 hit

from NCBI via the dataset and dataformat functions ³⁰. Hit genomes were retained only if complete, and

217 assembled genomes were used for further analysis if the alignment length was > 70% of the complete

218 hit geno 218 bit genome, indicating the assembly of a near-complete genome from wastewater (Figure S4b).

219 Scaffolds representing near-complete genomes for JC polyomavirus were collected for phylogenetic

220 analyses. Potential 218 hit genome, indicating the assembly of a near-complete genome from wastewater (**Figure S4b**).

219 Scaffolds representing near-complete genomes for JC polyomavirus were collected for phylogenetic

220 analyses. Potent 220 analyses. Potential assembly errors were inspected by mapping reads to assembled scaffolds using

221 bowtie2 and visualizing with the Integrative Genomics Viewer (IGV)³⁹. No assembly errors were

222 detected, and 221 bowtie2 and visualizing with the Integrative Genomics Viewer (IGV) ³⁹. No assembly errors were

222 detected, and representative mappings are shown in **Figure S5**. Given the circularity of the JC

223 polyomavirus g bowtie2 and visualizing with the Integrative Genomics Viewer (IGV) ³⁹. No assembly errors were

222 detected, and representative mappings are shown in **Figure S5**. Given the circularity of the JC

223 polyomavirus genom 222 detected, and representative mappings are shown in Figure 33. Given the chedularly of the Je
223 polyomavirus genome, assemblies were also examined in Geneious ⁴⁰, and repeated regions at the
224 beginning and the e polyomavirus genome, assemblies were also examined in Geneious ∞ , and repeated regions at the beginning and the end of the sequences were trimmed before the alignment (**Figure S6**). Multiple sequence alignment was per Example and the end of the sequences were trimmed before the alignment (Figure 36). Multiple sequence alignment was performed by MUSCLE ⁴¹ with all trimmed scaffolds, all JC polyomavirus reference genomes from NCBI GenB sequence alignment was performed by MUSCLE 44 with all trimmed scaffolds, all JC polyomavirus

226 reference genomes from NCBI GenBank released within two years (n=39), and the best-hit results from

227 BLASTn for each 227 BLASTn for each scaffold. The alignment was inspected in Geneious to identify a common starting point,
228 and all 10 scaffolds were recircularized to this point. The recircularized scaffolds were queried against
229 228 and all 10 scaffolds were recircularized to this point. The recircularized scaffolds were queried against
229 NCBI again to identify new best-hit reference genomes, which may have changed due to genome
230 curation. T 229 NCBI again to identify new best-hit reference genomes, which may have changed due to genome

230 curation. The final dataset included these curated genomes, new best-hits, and the 39 JC polyomavirus

231 references. A 220 Curation. The final dataset included these curated genomes, new best-hits, and the 39 JC polyomavirus
231 references. Alignment was performed with MUSCLE followed by GBlocks⁴² to identify informative
232 references. 231 references. Alignment was performed with MUSCLE followed by GBlocks⁴² to identify informative
232 regions, and MEGA 11.0⁴³ was used to generate and visualize the final maximum likelihood tree using
233 the Tamura references. Alignment was performed with MUSCLE followed by GBlocks⁴² to identify informative

132 regions, and MEGA 11.0⁴³ was used to generate and visualize the final maximum likelihood tree using

133 the Tamura Ne regions, and MEGA 11.0⁻³ was used to generate and visualize the final maximum likelihood tree using
233 the Tamura Nei model with 100 bootstrap replicates.
234 **2.6 Statistical analysis and data availability**

234 **2.6 Statistical analysis and data availability**
234 **2.6 Statistical analysis and data availability** 234 2.6 Statistical analysis and data availability

236 concentration and extraction methods were evaluated using the Kruskal-Wallis test, followed by post-
237 hoc pairwise Dunn's test. All statistical tests were performed using the Python package scipy stats, and
238 sign 237 hoc pairwise Dunn's test. All statistical tests were performed using the Python package scipy.stats, and
238 significance was determined at a 95% confidence interval ($p < 0.05$). Sequencing data for this project for
2 239 this project has been deposited in the NCBI Sequence Read Archive (SRA) under accession number:
240 SUB13892842 and Bioproject ID: PRJNA1047067. The processed data, reproducible code, and the
241 analysis workflow are

2366 concentration and extraction methods were evaluated using the Kruskal-Wallis test, followed by post-followed by p

240 SUB13892842 and Bioproject ID: PRJNA1047067. The processed data, reproducible code, and the
241 analysis workflow are available at https://github.com/mj2770/Wastewater-virus-surveillance-.
242 3. Results and Discussion SUB13892842 and Bioproject ID: PRJNA1047067. The processed data, reproducible code, and the
241 analysis workflow are available at https://github.com/mj2770/Wastewater-virus-surveillance-.
242 3. Results and Discussion
243 242 3. Results and Discussion
243 In this study, wastewater influent was collected from a single WWTP on three dates, and virus-
244 concentrated and extracted by four methods: IP method (Innovaprep ultrafiltration of liqu 242 3. Results and Discussion
243 In this study, wastewater influent wa
244 concentrated and extracted by four
245 paired with a small-volume extrac
246 performed on total influent paired
247 volume direct extraction), and 244 Concentrated and extracted by four methods: IP method (Innovaprep ultrafiltration of liquid portion
245 paired with a small-volume extraction kit), NT method (Nanotrap beads-based affinity capture
246 performed on tota Extra concentrated and extracted by four methods: IP method (Innovaprep ultrafiltration of liquid portion
245 paired with a small-volume extraction kit), **NT** method (Nanotrap beads-based affinity capture
246 performed on paired with a small-volume extraction kit), **NT** method (Nanotrap beads-based affinity capture
performed on total influent paired with small-volume extraction kit), **PMG** method (Promega large
247 volume direct extraction) 247 volume direct extraction), and **Solids method** (centrifugation paired with small-volume extraction kit).
248 volume direct extraction), and **Solids method** (centrifugation paired with small-volume extraction kit).
248 247 volume direct extraction), and Solids method (centrifugation paired with small-volume extraction kit).
248 The resulting 36 samples (12 samples in biological triplicate) were processed using the virus surveillance
250 249 panel (VSP) from Illumina using probe-capture enrichment. Following the initial analysis, an outlier
250 sample was identified, indicating unsuccessful library preparation (see Methods), and this sample was
251 **3.1 S**

250 sample was identified, indicating unsuccessful library preparation (see Methods), and this sample was
251 excluded from all analyses.
252 **3.1 Sample quality and sequence data**
253 The DNA and RNA generated using the 251 excluded from all analyses.
252 **3.1 Sample quality and sequence data**
253 The DNA and RNA generated using the four methods differed in concentration (Kruskal-Wallis test p =
254 2E-6 and 7E-7, respectively), fragment 252 **3.1 Sample quality and so.
253 The DNA and RNA generate
254 2E-6 and 7E-7, respectively
255 Solids method consistently
256 RNA (Figure 1a), while the** 252 3.1 Sample quality and sequence data

253 The DNA and RNA generated using the fo

254 2E-6 and 7E-7, respectively), fragment size

255 Solids method consistently resulted in yie

256 RNA (Figure 1a), while the IP meth 254 2E-6 and 7E-7, respectively), fragment size distribution, and RNA integrity (ANOVA test p = 1E-13). The
255 Solids method consistently resulted in yields that were higher than other methods for both DNA and
256 RNA (F 255 Solids method consistently resulted in yields that were higher than other methods for both DNA and RNA (Figure 1a), while the IP method, which includes a solids removal step, resulted in significantly lower total DNA 255 Solids method consistently resulted in yields that were higher than other methods for both DNA and

256 RNA (Figure 1a), while the IP method, which includes a solids removal step, resulted in significantly

257 Iower 256 RNA (Figure 1a), while the IP method, which includes a solids removal step, resulted in significantly lower total DNA and RNA yield compared to Solids and PMG (Figure 1a, IP v.s. Solids DNA $p = 3E-7$, IP v.s. PMG RNA 257 lower total DNA and RNA yield compared to Solids and PNG (Figure 1a, IP v.s. Solids DNA p = 3.2-7, IP v.s. PMG RNA p = 3.004). All methods yielded a higher concentration of RNA than DNA, but the resulting ratios of RN 259 higher concentration of RNA than DNA, but the resulting ratios of RNA:DNA varied significantly (Kruskal-
260 Wallis test $p = 0.002$) across methods from 2.0 ± 0.7 (for NT) to 4.3 ± 1.6 (for PMG). Unlike the other
261 265 annegrity is not discernible from the RNA Integrity Number (RIN) alone, the highest RIN was
266 observed with the PMG method (6.4 ± 1.0, Figure 1c), which suggested more intact prokaryotic RNA was 261 methods, shorter RNA fragments were observed with the NT method and 16S rRNA and 23S rRNA were
262 absent, perhaps accounting for the low RNA:DNA ratio. The lack of ribosomal RNA may be due to the
263 exclusion of bact 262 absent, perhaps accounting for the low RNA:DNA ratio. The lack of ribosomal RNA may be due to the exclusion of bacteria by the nanotrap hydrogel particle shells, which have specific pore sizes and are chemically modif 263 exclusion of bacteria by the nanotrap hydrogel particle shells, which have specific pore sizes and are
264 chemically modified to prevent the entry and capture of large or non-targeted particles ⁴⁴. Although viral
2 264 chemically modified to prevent the entry and capture of large or non-targeted particles ⁴⁴. Although viral
265 RNA integrity is not discernible from the RNA Integrity Number (RIN) alone, the highest RIN was
266 obser chemically modified to prevent the entry and capture of large or non-targeted particles ⁴⁴. Although viral
265 . RNA integrity is not discernible from the RNA Integrity Number (RIN) alone, the highest RIN was
266 observe

preserved with the PMG method.

268 After sequencing 36 samples, a total of 535 million reads were generated, averaging 14.86 ± 4.46 million

269 reads per sample (Figure 1b), and the removal of PCR duplicates reduced read 267 observed with the PMG method (6.4 ± 1.0, Figure 1c), which suggested more intact prokaryotic RNA was
267 preserved with the PMG method.
268 After sequencing 36 samples, a total of 535 million reads were generated, ave 268 After sequencing 36 samples, a to
269 reads per sample (Figure 1b), and
270 samples. As the IP method pro
271 surprising that after deduplication
272 1.3 million, Figure 1b) compared t
273 Solids p = 0.04). Nonetheless 269 reads per sample (**Figure 1b**), and the removal of PCR duplicates reduced read counts by over 50% for all

270 samples. As the IP method produced the lowest RNA and DNA input concentrations, it was not

271 surprising Eads per sample (Figure 1b), and the removal of PCR duplicates reduced read counts by over 50% for all
270 samples. As the IP method produced the lowest RNA and DNA input concentrations, it was not
271 surprising that aft 271 surprising that after deduplication these samples also retained significantly fewer unique reads (3.3 ± 1.3 million, **Figure 1b**) compared to samples from the Solids and NT methods (IP vs. NT p = 0.005, IP vs. Solids 272 1.3 million, **Figure 1b**) compared to samples from the Solids and NT methods (IP vs. NT $p = 0.005$, IP vs.
273 Solids $p = 0.04$). Nonetheless, the count of unique reads was not clearly related to the DNA and RNA
274 c 272 1.3 million, Figure 1b) compared to samples from the Solids and NT methods (iP vs. NT p = 0.005, iP vs.
273 Solids $p = 0.04$). Nonetheless, the count of unique reads was not clearly related to the DNA and RNA
274 conc 274 concentrations, perhaps due to the dilution of nucleic acids (**Table S1**) before library preparation, and
275 the multiple amplification and equimolar pooling steps during library preparation.
The multiple amplificatio 274 concentrations, perhaps due to the dilution of nucleic acids (Table S1) before library preparation, and
275 the multiple amplification and equimolar pooling steps during library preparation. 275 the multiple amplification and equimolar pooling steps during library preparation.

276

292 methods employed.

277
278
279
280
281
282 277 Figure 1. Nucleic acids and unique read counts by sample processing method. (a) Averaged concentrations of extracted DNA and RNA produced by each method (n=9 samples per method); (b) Averaged raw read counts and counts 279 unique reads after QC trimming and deduplication in each method (n = 9 samples for IP, NT, and Solids, n = 8 for PMG); (c)
280 Representative RNA fragment size distribution and average RNA integrity number (RIN) for ea

279 Unique reads after QC trimming and deduplication in each method (n = 9 samples for it, Nt, and Solids, n = 8 for ENN); Representative RNA fragment size distribution and average RNA integrity number (RIN) for each metho 281 Representative RNA fragment size distribution and average RNA integrity number (RIN) for each method. Note that samples
281 are diluted before fragment analysis (IP: undiluted, NT: 25x, PMG: 25x, Solids: 200x), so y-ax 282 **3.2 Taxonomic classification and virus composition similarity**
282 **3.2 Taxonomic classification and virus composition similarity**
284 the selected Minimum Hit Length (MHL) across all methods, and most classified rea 282 3.2 Taxonomic classification and virus composition similarity
283 0ver 40% of unique reads were not taxonomically classified by Re
284 the selected Minimum Hit Length (MHL) across all methods, and mothods the domain B 284 the selected Minimum Hit Length (MHL) across all methods, and most classified reads were assigned to
285 the domain Bacteria (ranging from 25.84 ± 6.81% to 40.88 ± 13.13%, **Figure 2a**). It is likely that a larger
286 285 the domain Bacteria (ranging from 25.84 \pm 6.81% to 40.88 \pm 13.13%, **Figure 2a**). It is likely that a larger proportion of unique reads would have received an assigned taxonomy at a lower classification stringenc 285 the domain Bacteria (ranging from 25.84 ± 6.81% to 40.88 ± 13.13%, Figure 2a). It is likely that a larger
286 proportion of unique reads would have received an assigned taxonomy at a lower classification
287 the downst 287 stringency; however, such low-confidence assignments have the potential to introduce substantial noise
288 to downstream assessments. Future functionalization of these platforms will require tuning of these
289 stringe 288 to downstream assessments. Future functionalization of these platforms will require tuning of these stringency thresholds for the desired application, balancing classification sensitivity with assignment confidence. Th 289 stringency thresholds for the desired application, balancing classification sensitivity with assignment
290 confidence. These findings could also reflect the current limitations of reference-based classifiers⁴⁶ and
2 290 confidence. These findings could also reflect the current limitations of reference-based classifiers⁴⁶ and
291 limited enrichment of targets using probe-capture, irrespective of the concentration and extraction
292 m confidence. These findings could also reflect the current limitations of reference-based classifiers⁴⁶ and
291 alimited enrichment of targets using probe-capture, irrespective of the concentration and extraction
292 meth 292 methods employed.
292 methods employed.

293

294
295
296 Figure 2. Taxonomic profiles of reads and virus hosts differed by interiod. (a) Domain-level classification of
295 Innique reads by Recentrifuge, with samples collected on three sampling dates and processed by four methods 295 unique reads by Recentringe, with samples conceted on three sampling dates and processed by four methods (n=3, except
296 Promega 4/26). "unclassified" is the sum of reads discarded by Recentrifuge without taxonomic cl 297 Promega 4/26). "unclassified" is the sum of reads discarded by Recentringe without taxonome classification and those
297 classified as "Root" but without a domain-level classification. "Human" represented unique reads 297 classified as "Root" but without a domain-level classification. Thailian Tepresented unique reads inapped to two downloaded
298 human genomes (see methods); (b) Percentages of unique reads identified as RNA, double-str reads in the manningenomes (see methods); (b) Percentages of unique reads identified as RNA, double-strand DNA, and single-strand DNA
299 buriuses based on kingdom-level virus classification; (c) Percentages of unique read viruses based on kingdom-level virus classification; (c) Percentages of unique reads identified as virus species linked to human

200 and non-human hosts in NCBI or for which species-level taxonomy was not determined; (d) 301 and non-human hosts in NCBI or for which species-level taxonomy was not determined; (d) refeemages of unique viral reads
301 associated with different host categories in the NCBI Virus database. Note that "human" in (c 302 Thuman & vertebrates" and "human" in (d). In (d), reads assigned to BCoV were subtracted from counts of reads assigned to 303 Thuman & vertebrates" and are not displayed. 303 thuman & vertebrates" and are not displayed.
304 The percentage of reads classified as viral ranged from 0.17 ± 0.02% (Solids) to 1.82 ± 0.46% (IP) of Promega 4/26). "unclassified" is the sum of reads discarded by Recentrifuge without taxonomic classification and those classified as "Root" but without a domain-level classification. "Human" represented unique reads mapped to two downloaded n o

305 unique reads across different methods (Figure 2b), surpassing the reported < 0.011% in untargeted 305 Unique reads across different methods (Figure 2b), surpassing the reported < 0.011% in untargeted sequencing ⁹. The IP samples yielded significantly higher percentages of viral reads than Solids and NT 305 unique reads across different methods (Figure 2b), surpassing the reported \leq 0.011% in untargeted sequencing ⁹. The IP samples yielded significantly higher percentages of viral reads than Solids and NT (1.82 ± 0 306 sequencing °. The IP samples yielded significantly higher percentages of viral reads than Solids and NT
307 (1.82 ± 0.46%, **Figure 2b**, IP vs. Solids p = 8E-7, and IP vs. NT p = 0.004), followed by the PMG samples 307 (1.82 ± 0.46%, Figure 2b, IP vs. Solids p = 8E-7, and IP vs. NT p = 0.004), followed by the PMG samples $\frac{1}{1}$

308 309 significantly more RNA viruses (Figure 2b) and viruses associated with human and/or vertebrate hosts
310 than NT and Solids methods (0.64 ± 0.27% human viruses in total unique reads from IP, Figures 2c and 309 significantly more RNA viruses (Figure 2b) and viruses associated with human and/or vertebrate hosts
310 than NT and Solids methods (0.64 ± 0.27% human viruses in total unique reads from IP, Figures 2c and
311 2d, IP d
I than NT and Solids methods ($0.64 \pm 0.27\%$ human viruses in total unique reads from IP, Figures 2c and 310 than NT and Solids methods (0.64 ± 0.27% human viruses in total unique reads from II, Figures 2c and
311 2d, IP vs. NT p = 0.002, IP vs. Solids p =1E-6). The IP and PMG methods incorporated a solids removal
312 step a 311 2d, IP vs. NT p = 0.002, IP vs. Solids p =1E-6). The IP and PMG methods incorporated a solids remova $\frac{1}{2}$ step after attempting to release solid-associated viruses by adding 5% Tween 20³⁵ or protease,
313 Frespectively ⁴⁶. These steps not only prevent clogging during sample processing but also strike a balance
314 Freshwee $\frac{1}{2}$ respectively ³⁶. These steps not only prevent clogging during sample processing but also strike a balance
314 . between eliminating solid-associated non-viral microorganisms like bacteria and attempting to retain
315 . v 314 between eliminating solid-associated non-viral microorganisms like bacteria and attempting to retain
315 viruses. As a result, a notably lower ratio of classified bacterial reads to classified viral reads was
316 obse 316 observed in IP and PMG samples (25 ± 14:1 and 38 ± 24:1, respectively) in comparison to Solids and NT
317 samples (241 ± 83 and 66 ± 12, respectively) (IP vs. NT p = 0.04; IP vs. Solids p =1E-5; PMG vs. Solids p = $\frac{1}{2}$ observed in IP and PMG samples (25 \pm 14:1 and 38 \pm 24:1, respectively) in comparison to Solids and NT 317 samples (241 ± 83 and 66 ± 12, respectively) (IP vs. NT $p = 0.04$; IP vs. Solids $p = 1E-5$; PMG vs. Solids $p = 318$ 0.0006). In NT and Solids samples, most viral reads were associated with bacterial hosts, based on th 318 0.0006). In NT and Solids samples, most viral reads were associated with bacterial hosts, based on the
319 NCBI taxonomy database (**Figure 2d**). This finding is consistent with the high fraction of DNA viruses in e
n 319 0.0006). In the same samples, most change as a second term in adventual method with the second of DNA viruses in
320 those samples (Figure 2b), as most bacteriophages are DNA viruses ⁴⁷. er
n 320 Thus taxonomy database (Figure 2d). This miding is consistent with the high fraction of DNA viruses in those samples (Figure 2b), as most bacteriophages are DNA viruses 47 .
 321 To compare virus composition acros n
V

those samples (**Figure 2b**), as most bacteriophages are DNA viruses ⁻⁻'.
321 . To compare virus composition across the four sample preparation me
322 . Recentrifuge were extracted from each sample, and MASH ³² wa 322 Recentrifuge were extracted from each sample, and MASH ³² was used to assess pairwise sequence
323 Similarity. In a principal component analysis using MASH distances, triplicate samples clustered together Recentrifuge were extracted from each sample, and MASH ²⁴ was used to assess pairwise sequence

S23 similarity. In a principal component analysis using MASH distances, triplicate samples clustered together

S24 (PERMANOV 323 similarity. In a principal component analysis using MASH distances, triplicate samples clustered together
324 (PERMANOVA test $p = 0.985$), while all samples were separated by concentration/extraction methods
325 along 325 along PC1 (37.2% of the variation, **Figure 3**, PERMANOVA test $p = 0.001$). Specifically, IP- and PMG 326 samples clustered together, while NT and Solids samples were distinct (**Figure 3**). The predominance of bacterio 325 along PC1 (37.2% of the variation, Figure 3, PERMANOVA test p = 0.001). Specifically, IT- and PMG
326 samples clustered together, while NT and Solids samples were distinct (Figure 3). The predominance of
327 bacteri 326 samples clustered together, while NT and Solids samples were distinct (Figure 3). The predominance of
328 bacteriophage in both NT and Solids samples likely contributed to their differentiation from the other
328 two m 328 two methods. Samples were separated by sampling dates along PC2 (24.5% of the variation, **Figure 3,**
329 PERMANOVA test p = 0.001), with samples from March 1, 2023, differing from those collected on April 328 two methods. Samples were separated by sampling dates along PC2 (24.5% of the variation, Figure 3,
329 PERMANOVA test $p = 0.001$), with samples from March 1, 2023, differing from those collected on April
330 19 and Ap 329 PERMANOVA test p = 0.001), with samples from March 1, 2023, differing from those collected on April
330 19 and April 26. This differentiation was observed consistently across all four methods. These temporal
331 shift 331 shifts in virus composition may suggest a temporally variable metavirome composition in wastewater
332 potentially influenced by changes in circulating viruses ^{8, 48, 49} and changing wastewater conditions, such 332 potentially influenced by changes in circulating viruses ^{8, 48, 49} and changing wastewater conditions, such
333 as flow rate, total suspended solids (TSS), total organic compounds (TOC), and the abundance of 332 potentially influenced by changes in circulating viruses of the thanging wastewater conditions, such
333 as flow rate, total suspended solids (TSS), total organic compounds (TOC), and the abundance of
334 antagonistic 334 antagonistic microorganisms 50,51 . 334 antagonistic microorganisms $^{50, 51}$.
 0.15 $\frac{1}{2}$ ີງ
f -
er:
r
İ ;
|
| $, |$ $\frac{1}{2}$ ์
|
| י
f

335
336
337

1 Figure 3. Viral sequence composition was immediced by wastewater virus concentration/extraction
337 method and sample date. Principal component analysis (PCoA) plot was generated using the MASH distance, which was 337 method and sample date. Principal component analysis (PCoA) plot was generated using the MASH distance, which was s

338 calculated based on sequence similarity among all reads classified 339 colors, and different sampling dates are represented by shapes.
340 **3.3 Human virus species richness and composit** calculated based on sequence similarity among all reads classified as viral by Centrifuge. Different methods are represented by

340 **3.3 Human virus species richness and composition**
341 PMG and IP methods yielded higher species-level richne 341 PMG and IP methods yielded higher species-level richness of total viruses detected with >10 reads (241
342 and 176 viruses, respectively) and human viruses (20 and 26 respectively) compared to NT and Solids 342 and 176 viruses, respectively) and human viruses (20 and 26 respectively) compared to NT and Solids
343 (Figure S3a), although total read depth was similar for all samples (Figure 1b, p = 0.44). Thus, removing
344 sol 343 (Figure S3a), although total read depth was similar for all samples (Figure 1b, p = 0.44). Thus, removing
344 solids after releasing solid-associated viruses did not compromise the richness of detected human 344 solids after releasing solid-associated viruses did not compromise the richness of detected human
345 viruses. Conversely, including solids produced lower species-level diversity. Of the 66 virus "groups" of
346 high p 345 viruses. Conversely, including solids produced lower species-level diversity. Of the 66 virus "groups" of
346 high public health significance listed as targets in the Illumina VSP panel (Table S5), IP samples detected 346 high public health significance listed as targets in the Illumina VSP panel (Table S5), IP samples detected
347 members of 11 (Figure S3a). These included human coronavirus-OC43 (hCoV-OC43), adenovirus, 347 high public health significance listed as targets in the Illumina VSP panel (Table 35), it samples detected
348 astrovirus, aichivirus, enterovirus, norovirus, coxsackievirus, rotavirus, salivirus, and sapovirus, as we members of 11 (Figure S3a). These included human coronavirus-OC43 (hCoV-OC43), adenovirus, as well as
348 astrovirus, aichivirus, enterovirus, norovirus, coxsackievirus, rotavirus, salivirus, and sapovirus, as well as
349 349 astronomy members, and the exact list of species and strains used by Illumina for probe design is
350 proprietary; we note that enteroviruses are a diverse group which contains coxsackieviruses, while 349 mpox (Figure S3b), though the exact list of species and strains used by indifficult probe design is
350 proprietary; we note that enteroviruses are a diverse group which contains coxsackieviruses, while
351 hCoV-OC43 i $\frac{1}{s}$ e
S
C
S on
of ^{of}
d $\frac{1}{2}$ $\frac{5}{5}$ s
e e

- 351 hCoV-OC43 is a sub-species level category.

Proprietary are a diverse group which contains c P
351 Rotavirus A
351 hCovernative level category. $2R$ 3468 Severe acute respiratory syndrome-related coronavirus Sapporo virus -Norwalk virus 3663 1648 4284 434 Enterovirus A Enterovirus B 68 87 Enterovirus C $\overline{17}$ Human enterovirus (unclassified) Aichivirus A Percentage mean Mamastrovirus 1 2003 9156 7244 695 1061 Human astrovirus (unclassified) 1575 $\overline{104}$ $1e-1$ Mamastrovirus sp. (unclassified) -1667 224 731 Kobuvirus sp. (unclassified) $1e-2$ **RNA virus** Picobirnavirus sp. (unclassified) Betapolyomavirus hominis 1299 leq -3 182 Betapolyomavirus secuhominis 1044 1259 224 176 Alphapolyomavirus quintihominis Betapolyomavirus quartihominis Human mastadenovirus A Human mastadenovirus B Human mastadenovirus C Human mastadenovirus D 117 Human mastadenovirus E Human mastadenovirus F Monkeypox virus -Alphapapillomavirus 6 Alphapapillomavirus 7 **DNA** virus **AI26 A** 2^6 413 介 0119 介 介 **AI** 29 A76 ŵ **0119** 126 date
-

352

353
354 354 Frelative abundance of each virus species in total unique reads across triplicate samples, based on Recentrifuge
1955 Fread classification. Species with fewer than an average of 10 reads per sample are not shown. Text Feative abundance of each virus species in total unique reads across triplicate samples, based on Recentrifuge
355 read classification. Species with fewer than an average of 10 reads per sample are not shown. Text in each S35 read classification. Species with fewer than an average of 10 reads per sample are not shown. Text in each cent
356 indicates the average read counts assigned to the species for each sample. Viruses are grouped by geno S336 Indicates the average read counts assigned to the species for each sample. Viruses are grouped by genome type
357 INCBI taxIDs corresponding to species without names (e.g. "sp.") are appended with "(unclassified)" (se 357 NCBI taxIDs corresponding to species without names (e.g. "sp.") are appended with "(
358 supplementary methods). Note that Betacoronavirus 1 includes the spike-in bovine coronavirus.
359 358 supplementary methods). Note that Betacoronavirus 1 includes the spike-in bovine coronavirus. Figure 4. Relative abundance of human virus species in each sample. Fill indicates the average percent . e

361 the four methods (**Figure 4**). Some viruses were consistently detected by all methods, including human polyomavirus, mastadenovirus, mamastrovirus 1, and norwalk virus, which are known to be shed at high concentrations 362 polyomavirus, mastadenovirus, mamastrovirus 1, and norwalk virus, which are known to be shed at high concentrations in human waste 5, 9, 10, 13, 23, 48, 52-55. RNA virus species, including severe acute respiratory synd 363 concentrations in human waste $^{5, 9, 10, 13, 23, 48, 52\cdot 55}$. RNA virus species, including severe acute respiratory syndrome-related coronavirus, sapporo virus, and enteroviruses were not detected in NT and Solids concentrations in human waste $^{5, 9, 10, 13, 23, 48, 52-55}$. RNA virus species, including severe acute respiratory
syndrome-related coronavirus, sapporo virus, and enteroviruses were not detected in NT and Solids
sample 369 (e.g., Dengue, Chikungunya), bloodborne viruses (e.g., Hepatitis virus and HIV), or hemorrhagic fever-
370 related viruses (e.g., lassa mamarenavirus, junin virus, etc.) were detected, despite their inclusion in the
37 366 human mastadenovirus B, D, and F were detected in all samples, while human mastadenovirus A, C, and
367 E were not detected in certain samples (**Figure 4**). This variability suggested that related virus species
368 ma E were not detected in certain samples (**Figure 4**). This variability suggested that related virus species
368 may be differentially detected by different concentration methods. No arthropod-transmitted viruses
369 (e.g., E were not detected in certain samples (Figure 4). This variability suggested that related virus species
368 may be differentially detected by different concentration methods. No arthropod-transmitted viruses
369 (e.g., De (e.g., Dengue, Chikungunya), bloodborne viruses (e.g., Hepatitis virus and HIV), or hemorrhagic fever-
370 felated viruses (e.g., lassa mamarenavirus, junin virus, etc.) were detected, despite their inclusion in the
371 pr 371 probe panel. Mpox, detected intermittently in wastewater since the outbreak in 2022 ^{56, 57}, was detected
372 at low levels in IP, PMG, and NT samples. Overall, it seems reasonable to conclude that these results
373 g probe panel. Mpox, detected intermittently in wastewater since the outbreak in 2022 ^{36, 37}, was detected

at low levels in IP, PMG, and NT samples. Overall, it seems reasonable to conclude that these results

generally r

375 3.4 Potential of recovering near-complete human virus genomes

373 at Potential of recovering near-complete human virus genomes
374 Bay Area at the time of sample collection.
375 **3.4 Potential of recovering near-complete human virus genomes**
376 Seven near-complete human virus genome 374 Bay Area at the time of sample collection.

375 3.4 Potential of recovering near-complete human virus genomes

376 Seven near-complete human virus genomes were assembled from IP samples, the most from any

377 concent 375 3.4 Potential of recovering near-comp
376 Seven near-complete human virus geno
377 concentration/extraction method (Figure
378 human virus reads in these samples (59,9
379 near-complete human virus genomes wer 375 3.4 Potential of recovering near-complete human virus genomes
376 Seven near-complete human virus genomes were assembled from II
377 concentration/extraction method (Figure S4b). This aligned with the P
378 human viru 377 concentration/extraction method (Figure S4b). This aligned with the high numbers of total virus and
378 human virus reads in these samples (59,965 ± 28,180 and 20,242 ± 9,294, respectively, Table S1). No
379 near-comp Extraction method (Figure 34b). This aligned with the high numbers of total virus and
378 human virus reads in these samples (59,965 ± 28,180 and 20,242 ± 9,294, respectively, **Table S1**). No
379 near-complete human virus maniar virus reads in these samples (59,965 ± 28,180 and 20,242 ± 9,294, respectively, rabic 31, increase near-complete human virus genomes were obtained from Solids-extracted samples (**Figure S4b**) likely due to insuffici The to insufficient reads for total viruses and human viruses (11,043 ± 2,720 and 213 ± 99, respectively,
381 due to insufficient reads for total viruses and human viruses (11,043 ± 2,720 and 213 ± 99, respectively,
381 T

381 **Table S1**). These results highlight the need to understand the minimum sequencing depth in relation to the proportion of viral reads required for the assembly of high-quality virus genomes.
383 IC polyomavirus compos 382 Table S1). These results highlight the need to understand the minimum sequencing depth in relation to
382 IC polyomavirus composite genomes were assembled in samples from three concentration/extraction
384 IC polyomavi 383 IC polyomavirus composite genomes were assembled in samples from three concernent methods (IP, PMG, and NT) and multiple replicates (**Figure S4b**). The recovery of genomes is perhaps unsurprising given that approximate 384 methods (IP, PMG, and NT) and multiple replicates (**Figure S4b**). The recovery of JC polyomavirus genomes is perhaps unsurprising given that approximately 40% of the population sheds the virus through urine ⁵⁴. Also, 385 genomes is perhaps unsurprising given that approximately 40% of the population sheds the virus
386 through urine ⁵⁴. Also, as a non-enveloped DNA virus with a circular genome, JC polyomavirus is highly
387 resistant 386 through urine ⁵⁴. Also, as a non-enveloped DNA virus with a circular genome, JC polyomavirus is highly
387 resistant to environmental stress and exonuclease activity ⁹. Ten scaffolds classified as near-complete JC through urine ⁵⁴. Also, as a non-enveloped DNA virus with a circular genome, JC polyomavirus is highly

resistant to environmental stress and exonuclease activity⁹. Ten scaffolds classified as near-complete JC

polyom resistant to environmental stress and exonuclease activity $\%$. Ten scaffolds classified as near-complete JC
polyomavirus genomes were used for phylogenetic analysis. At least one subtype of JC polyomavirus 3
389 was pre was present (Node 1353 NT_301_1), affiliated with clades from South Africa (**Figure S7**). Although other
scaffolds were clustered together, they exhibited relatively low node support values (< 50); likely several
of these was present (Node 1333 NT_301_1), armated with clades from South Africa (Figure 37). Although other
scaffolds were clustered together, they exhibited relatively low node support values (< 50); likely several
of these scaff 391 of these scaffolds represent the same JC polyomavirus population in replicate wastewater samples, with
392 variations in the composite assembly. These results, and those from other recent studies 25 ,
393 demonstra 392 variations in the composite assembly. These results, and those from other recent studies 25 , demonstrated that probe capture enrichment can yield whole genomes of high-abundance viruses for phylogenetic analysis, variations in the composite assembly. These results, and those from other recent studies $\frac{25}{7}$, demonstrated that probe capture enrichment can yield whole genomes of high-abundance viruses for phylogenetic analysis, w 394 phylogenetic analysis, which may be useful for identifying novel virus strains in the future.

Sensing ment can yield whole genomes for the future virus of the future of the future of the future of the strains in the f

394 phylogenetic analysis, which may be useful for identifying novel virus strains in the future.

397

398
399
400
402
403 Figure 5. Detection sensitivity comparison between dPCR and reads-based classification (Recentrifuge)
399 of sequencing results. (a) SARS-CoV-2 detection comparison; (b) BCoV detection comparison. Blue bars on the left y-a 399 of sequencing results. (a) SARS-CoV-2 detection comparison, (b) BCoV detection comparison. Blue Bars on the left y-axis
399 represent the virus concentration measured by dPCR in the final total nucleic acids (TNA) elut

For the virus concentration measured by discussional limit of detection are shown with a blue "x", and samples without
402 measurement were labeled with a black "x". Red points on the right y-axis represent virus read coun For Samples with dividend below the operational limit of detection are shown with a blue "x", and samples without measurement were labeled with a black "x". Red points on the right y-axis represent virus read counts from u discussion and the labeled with a black "x". Red points on the right y-axis represent virus read counts from dingle reads. The discussion discussion of the right of discussions of the right y-axis represent virus read coun 403 dashed red line at 10 reads indicates the operationia limit of detection used elsewhere in the analysis.
405 BCoV were quantified in the final extracted nucleic acids produced by each conce
406 sequencing, both SARS-Co BCoV were quantified in the final extracted nucleic acids produced by each concentration method. By
406 sequencing, both SARS-CoV-2 and BCoV were detected in PMG and IP samples at the employed
407 alignment stringency and Equivalently, both SARS-CoV-2 and BCoV were detected in PMG and IP samples at the employed alignment stringency and read count threshold (> 10 reads, **Figure 5**), which corresponded with the higher relative abundances o alignment stringency and read count threshold (> 10 reads, **Figure 5**), which corresponded with the higher relative abundances of human viruses in these two methods. However, the absolute concentrations of SARS-CoV-2 were and the method of state of the method (> 10 reads, Figure 5), which corresponded with the
A08 higher relative abundances of human viruses in these two methods. However, the absolute
and concentrations of SARS-CoV-2 were s 409 concentrations of SARS-CoV-2 were significantly lower in IP samples than in PMG samples (IP v.s. PMG p
410 = 0.009, **Figure 5a** and **Table S6**). Target virus concentrations could be increased by increasing the
 $\frac{1}{2$ 410 = 0.009, Figure 5a and Table S6). Target virus concentrations could be increased by increasing the $\frac{1}{2}$ 410 = 0.009, Figure 5a and Table S6). Target virus concentrations could be increased by increasing the

411 effective volume of wastewater processed. Specifically, the final volume of the ultrafilter concentrate
412 nearly always exceeded the maximum input for nucleic acid extraction, resulting in a lower effective
413 volu wet solids due to the limitation of the extraction kit, resulting in a lower effective sample volume

processed relative to the PMG and NT methods. The limited input may partially explain the low

concentrations of target volume (ranging from 16.3 ± 13 mL, Table S1). Similarly, only 0.25 g was extracted from 0.00 ± 0.18 g of
414 wet solids due to the limitation of the extraction kit, resulting in a lower effective sample volume
415 concent Frame is a processed relative to the PMG and NT methods. The limited input may partially explain the low

416 concentrations of targets observed by dPCR. Notably, although samples from March 1 showed similar

417 SARS-CoV concentrations of targets observed by dPCR. Notably, although samples from March 1 showed similar
417 SARS-CoV-2 concentrations from both NT and PMG methods (NT: 14.5 ± 3.4 gc/ μ L, PMG: 9.2 ± 1.4 gc/ μ L,
418 p = 0 SARS-CoV-2 concentrations from both NT and PMG methods (NT: 14.5 ± 3.4 gc/ μ L, PMG: 9.2 ± 1.4 gc/ μ L,
 $p = 0.18$, Table S6), no SARS-CoV-2 reads were detected in the NT samples from this date. Meanwhile,

although 418 p = 0.18, Table S6), no SARS-CoV-2 reads were detected in the NT samples from this date. Meanwhile,
419 although BCoV was detected by dPCR in NT and Solids samples at low levels, it was absent in the
420 sequencing res 418 p = 0.18, Table 36), no SARS-COV-2 reads were detected in the NT samples from this date. Meanwhile,
419 although BCoV was detected by dPCR in NT and Solids samples at low levels, it was absent in the
420 sequencing res

420 sequencing results. This suggests that in addition to the absolute viral concentration indicated by dPCR,
421 background non-target sequences may also influence target detection by sequencing.
422 **3.6 Implications for** 121 background non-target sequences may also influence target detection by sequencing.

122 3.6 Implications for genome surveillance of known and novel human viruses

123 Based on the comparisons reported here, wastewater 422 **3.6 Implications for genome surveillance of known and novel human viruses**
423 Based on the comparisons reported here, wastewater virus concentration/extraction r
424 chosen carefully and aligned with the specific mon 422 3.6 Implications for genome surveillance of known and novel human viruses
423 Based on the comparisons reported here, wastewater virus concentration/extraction
424 chosen carefully and aligned with the specific monitor 424 chosen carefully and aligned with the specific monitoring endpoint and goal (e.g., sequencing or dPCR,
425 specific targets or broad range of targets). Removing wastewater solids, after treatment with either
426 Tween 425 specific targets or broad range of targets). Removing wastewater solids, after treatment with either
426 Tween 20 (IP method) or protease (PMG method) and prior to concentration and extraction, improved
427 the overal Tween 20 (IP method) or protease (PMG method) and prior to concentration and extraction, improved

427 the overall detection of human viruses in probe-capture sequencing by minimizing the ratio between

428 off-target seq the overall detection of human viruses in probe-capture sequencing by minimizing the ratio between
428 off-target sequences and targeted human virus sequences 58 (Figure 2). However, solids removal may
429 also decrea off-target sequences and targeted human virus sequences ⁵⁸ (Figure 2). However, solids removal may
also decrease the sensitivity of virus detection by dPCR by decreasing the absolute quantities of the
target in the sampl off-target sequences and targeted human virus sequences ⁵⁸ (Figure 2). However, solids removal may
also decrease the sensitivity of virus detection by dPCR by decreasing the absolute quantities of the
target in the samp target in the sample (**Figure 5**)¹⁸. As untargeted sequencing was not performed in parallel, the extent to

431 which solids removal improved probe-capture enrichment specifically cannot be directly quantified.

432 Giv target in the sample (Figure 5) ²⁶. As untargeted sequencing was not performed in parallel, the extent to which solids removal improved probe-capture enrichment specifically cannot be directly quantified.
432 Given that 432 Given that methods that included solids showed higher relative abundances of DNA viruses, a DNase
433 treatment may improve the recovery of human RNA viruses with these methods. Additionally, while
434 only two extrac treatment may improve the recovery of human RNA viruses with these methods. Additionally, while

any two extraction methods were applied here (Qiagen AllPrep PowerViral and Promega Wizard Enviro

435 TNA), the extraction m 434 only two extraction methods were applied here (Qiagen AllPrep PowerViral and Promega Wizard Enviro
435 TNA), the extraction method used for solids and viral concentrates may affect the overall sensitivity of
436 seque TNA), the extraction method used for solids and viral concentrates may affect the overall sensitivity of
436 sequencing by influencing the degree of viral lysis and integrity of the resulting nucleic acids ¹⁶. Further
43 TNA), the extraction method used for solids and viral concentrates may affect the overall sensitivity of
436 sequencing by influencing the degree of viral lysis and integrity of the resulting nucleic acids 16 . Further sequencing by influencing the degree of viral lysis and integrity of the resulting nucleic acids ²⁶. Further

studies should be performed during periods of higher target concentration in wastewater (e.g., SARS-

CoV-2 su

443 included in the RVOP panel $^{14, 15, 49}$. However, in the present study and other studies using broad virus using different concentration and extraction methods.

440 Finally, the choice of probe panel likely also impacts the sensitivity of virus detection in probe-capture

441 sequencing. When using the RVOP probe set (which co Finally, the choice of probe panel likely also impacts t
441 sequencing. When using the RVOP probe set (which co
442 several studies found remarkably high coverages of SA
443 included in the RVOP panel ^{14, 15, 49}. Howeve Finally, the choice of probe set (which contains fewer virus targets than the VSP probe set)
several studies found remarkably high coverages of SARS-CoV-2, surpassing that of other human viruses
included in the RVOP panel several studies found remarkably high coverages of SARS-CoV-2, surpassing that of other human viruses
included in the RVOP panel $^{14, 15, 49}$. However, in the present study and other studies using broad virus
capture pa included in the RVOP panel $^{14, 15, 49}$. However, in the present study and other studies using broad virus capture panels $^{23, 48}$, sequence data were dominated by enteric viruses such as mamastrovirus, with limited d included in the RVOP panel $^{14, 15, 49}$. However, in the present study and other studies using broad virus capture panels $^{23, 48}$, sequence data were dominated by enteric viruses such as mamastrovirus, with limited d capture panels ^{23, 48}, sequence data were dominated by enteric viruses such as mamastrovirus, with
445 limited detection of SARS-CoV-2. This points to the inherent challenge of using broad panels as a means
446 of wastew 446 of wastewater-based surveillance for early detection of novel virus strains, which may appear at low
447 abundance before going on to cause a larger outbreak.
448 4. Acknowledgements
49 Funding was provided by the UCOP

abundance before going on to cause a larger outbreak.

448 4. Acknowledgements

449 Funding was provided by the UCOP Lab Fees CRT Award (L22CR4507). We thank Khi Lai at EBMUD for

450 sample collection and Sanaiya Islam fo 448 4. Acknowledgements
449 Funding was provided by the UCOP Lab Fees CRT Awa
450 sample collection and Sanaiya Islam for laboratory m
451 with advice from Justin Choi and Byran Bach at the Fun 448 4. Acknowledgements
449 Funding was provided by the U
450 sample collection and Sanaiya
451 with advice from Justin Choi and
451 450 sample collection and Sanaiya Islam for laboratory management. Library preparation was performed
451 with advice from Justin Choi and Byran Bach at the Functional Genomics Laboratory and sequencing was
451 with advice with advice from Justin Choi and Byran Bach at the Functional Genomics Laboratory and sequencing was performed. Library preparation was performed. Library preparation was performed. Library preparation was performed. The s 451 with advice from Justin Choi and Byran Bach at the Functional Genomics μ

453 thank Allie Nguyen and Van Trinh for their assistance with laboratory and bioinformatic analyses.
454 5. References
455 1. World Health, O., Guidelines for environmental surveillance of poliovirus circulation. In World 454 5. References
455 1. World Health, O., Guidelines for environmental surveillance of poliovirus circulation. In W
456 Health Organization: Geneva, 2003.
457 2. National Academies of Sciences, E.; Medicine; Health; Medic 454 5. References

455 1. World Healt

456 Health Organization:

457 2. National Aca

458 Board on Population

459 Wastewater-based II

460 *Action*, National Aca

Health Organization: Geneva, 2003.

456 Health Organization: Geneva, 2003.

457 2. National Academies of Sciences, E.; Medicine; Health; Medicine, D.; Division on, E.; Life, S.;

458 Board on Population, H.; Public Health,

457 2. National Academies of Scien
458 Board on Population, H.; Public Heal
459 Wastewater-based Infectious Diseas
460 *Action*, National Academies Press (U
461 Copyright 2023 by the National Acad
462 3. Boehm, A. B.; Hugh Board on Population, H.; Public Health, P.; Water, S.; Technology, B.; Committee on Community

459 Wastewater-based Infectious Disease, S., In Wastewater-based Disease Surveillance for Public Heal

460 Action, National Aca Wastewater-based Infectious Disease, S., In Wastewater-based Disease Surveillance for Public H.

460 Action, National Academies Press (US)

461 Copyright 2023 by the National Academy of Sciences. All rights reserved.: Wash Wastewater-based Infectious Disease, S., In Wastewater-based Disease Surveillance for Fublic Health
Action, National Academies Press (US)
462 3. Boehm, A. B.; Hughes, B.; Duong, D.; Chan-Herur, V.; Buchman, A.; Wolfe, M. K 460 Action, National Academies Press (03)
461 Copyright 2023 by the National Academ
462 3. Boehm, A. B.; Hughes, B.; Duor
463 Wastewater concentrations of human
465 study. *Lancet Microbe* 2023, 4, (5), e34
466 4. Nemudryi

462 3. Boehm, A. B.; Hughes, B.; Duong, D.; Chan-Herur, V.; Buchman, A.; Wolfe, M. K.; White, B.
463 Wastewater concentrations of human influenza, metapneumovirus, parainfluenza, respiratory syn
464 virus, rhinovirus, and Wastewater concentrations of human influenza, metapneumovirus, parainfluenza, respiratory syncyt

464 virus, rhinovirus, and seasonal coronavirus nucleic-acids during the COVID-19 pandemic: a surveilland

465 study. *Lance* Virus, rhinovirus, and seasonal coronavirus nucleic-acids during the COVID-19 pandemic: a surveillance

465 study. *Lancet Microbe* 2023, 4, (5), e340-e348.

466 4. Nemudryi, A.; Nemudraia, A.; Wiegand, T.; Surya, K.; Buyu 465 study. *Lancet Microbe* 2023, 4, (5), e340-e348.
466 4. Nemudryi, A.; Nemudraia, A.; Wiegand, T.; Surya, K.; Buyukyoruk, M.; Cicha, C.; Vanderwood, K
467 K.; Wilkinson, R.; Wiedenheft, B., Temporal Detection and Phylog 465 study. Lancet Microbe 2023, 4, (3), e340-e348.
466 4. Nemudryi, A.; Nemudraia, A.; Wiegand,
467 K.; Wilkinson, R.; Wiedenheft, B., Temporal Det
468 Municipal Wastewater. *Cell Rep Med* 2020, 1, (
469 5. Reyne, M. I.; A K.; Wilkinson, R.; Wiedenheft, B., Temporal Detection and Phylogenetic Assessment of SARS-CoV-2 in

468 Municipal Wastewater. *Cell Rep Med* 2020, 1, (6), 100098.

5. Reyne, M. I.; Allen, D. M.; Levickas, A.; Allingham, P.

Municipal Wastewater. Cell Rep Med 2020, 1, (6), 100098.

469 5. Reyne, M. I.; Allen, D. M.; Levickas, A.; Allingham, P.; Lock, J.; Fitzgerald, A.; McSparron, C.;

470 Nejad, B. F.; McKinley, J.; Lee, A.; Bell, S. H.; Quic

468 Municipal Wastewater. Cemep Med 2020, 1, (0), 100098.
469 5. Reyne, M. I.; Allen, D. M.; Levickas, A.; Allingham, F
470 Nejad, B. F.; McKinley, J.; Lee, A.; Bell, S. H.; Quick, J.; Hould
471 McGrath, J. W., Detection o Nejad, B. F.; McKinley, J.; Lee, A.; Bell, S. H.; Quick, J.; Houldcroft, C. J.; Bamford, C. G. G.; Gilpin, D.
471 McGrath, J. W., Detection of human adenovirus F41 in wastewater and its relationship to clinical ca
472 of a McGrath, J. W., Detection of human adenovirus F41 in wastewater and its relationship to clinical cases

472 of acute hepatitis of unknown aetiology. *Sci Total Environ* 2023, 857, (Pt 2), 159579.

473 6. Spurbeck, R. R.; M of acute hepatitis of unknown aetiology. *Sci Total Environ* **2023,** 857, (Pt 2), 159579.

473 6. Spurbeck, R. R.; Minard-Smith, A.; Catlin, L., Feasibility of neighborhood and building scale

474 wastewater-based genomic 472 of acute hepatitis of unknown aetiology: *Sci Total Environ 2023*, 857, (Pt 2₎, 159579.
473 6. Spurbeck, R. R.; Minard-Smith, A.; Catlin, L., Feasibility of neighborhood and
474 wastewater-based genomic epidemiology

wastewater-based genomic epidemiology for pathogen surveillance. Sci Total Environ 2021, 789,
475 147829.
476 7. Bisseux, M.; Debroas, D.; Mirand, A.; Archimbaud, C.; Peigue-Lafeuille, H.; Bailly, J. L.; Henc
477 C., Monit 474 wastewater-based genomic epidemiology for pathogen survemance. Scr Total Environ 2021, 789,
475 147829.
476 7. Bisseux, M.; Debroas, D.; Mirand, A.; Archimbaud, C.; Peigue-Lafeuille, H.; Bailly, J. L.; Her
477 C., Moni 476 7. F
477 C., Monit
478 complem
479 8. F
481 9. F
481 9. Figuerola

C., Monitoring of enterovirus diversity in wastewater by ultra-deep sequencing: An effective

478 complementary tool for clinical enterovirus surveillance. Water Res 2020, 169, 115246.

479 8. Brinkman, N. E.; Fout, G. S.; 477 C., Monitoring of enterovirus diversity in wastewater by ultra-deep sequencing: An effective

478 complementary tool for clinical enterovirus surveillance. Water Res 2020, 169, 115246.

479 8. Brinkman, N. E.; Fout, G. 478 Complementary tool for climical enterovirus surveillance. Water Res 2020, 109, 115246.
480 Seasonal Dynamics of Enterovirus Infections. *mSphere* 2017, 2, (3).
481 9. Fernandez-Cassi, X.; Timoneda, N.; Martinez-Puchol,

Seasonal Dynamics of Enterovirus Infections. *mSphere* 2017, 2, (3).

481 9. Fernandez-Cassi, X.; Timoneda, N.; Martinez-Puchol, S.; Rusinol, M.; Rodriguez-Manzano, J.;

482 Figuerola, N.; Bofill-Mas, S.; Abril, J. F.; Gir 480 Seasonal Dynamics of Enterovirus Infections. *Inspirent 2011, 2, (3)*.
481 9. Fernandez-Cassi, X.; Timoneda, N.; Martinez-Puchol, S.; Rus
482 Figuerola, N.; Bofill-Mas, S.; Abril, J. F.; Girones, R., Metagenomics f
483 Figuerola, N.; Bofill-Mas, S.; Abril, J. F.; Girones, R., Metagenomics for the study of viruses in urban

483 sewage as a tool for public health surveillance. *Sci Total Environ* **2018**, 618, 870-880.

484 10. Cantalupo, P Example 183 Sewage as a tool for public health surveillance. Sci Total Environ 2018, 618, 870-880.

484 10. Cantalupo, P. G.; Calgua, B.; Zhao, G.; Hundesa, A.; Wier, A. D.; Katz, J. P.; Grabe, M.; Hendr

485 W.; Girones, 483 sewage as a tool of public health surveillance. Scribton Environ 2018, 010, 070-880.
484 10. Cantalupo, P. G.; Calgua, B.; Zhao, G.; Hundesa, A.; Wier, A. D.; Katz, J. P.; Gra
485 W.; Girones, R.; Wang, D.; Pipas, J. M W.; Girones, R.; Wang, D.; Pipas, J. M., Raw Sewage Harbors Diverse Viral Populations. *mBio* 2011, 2, (5),
10.1128/mbio.00180-11.
13. Bibby, K.; Peccia, J., Identification of Viral Pathogen Diversity in Sewage Sludge by M

486 10.1128/mbio.00180-11.

486 10.1128/mbio.00180-11.

487 11. Bibby, K.; Peccia, J., Identification of Viral Pathogen Diversity in Sewage Sludge by Metagenome

488 Analysis. *Environmental Science & Technology* 2013, 487 11. Bibby, K.; Peccia,
488 Analysis. *Environmental S*
489 12. Gaudin, M.; Desn
490 to Human Infectious Disea
491 13. Martinez-Puchol,
492 Anton, A.; Abril, J. F.; Giro
493 NGS tools and occurrence Analysis. *Environmental Science & Technology* 2013, 47, (4), 1945-1951.

489 12. Gaudin, M.; Desnues, C., Hybrid Capture-Based Next Generation Sequencing and Its Application

490 to Human Infectious Diseases. *Front M* Analysis. Environmental Science & Analy, recimology 2013, 47, (4), 1945-1991.

489 12. Gaudin, M.; Desnues, C., Hybrid Capture-Based Next Generation Sequ

490 to Human Infectious Diseases. *Front Microbiol* **2018,** 9, 2924

490 to Human Infectious Diseases. *Front Microbiol* **2018**, 9, 2924.
491 13. Martinez-Puchol, S.; Rusinol, M.; Fernandez-Cassi, X.; Timoneda, N.; Itarte, M.; Andres, C.;
492 Anton, A.; Abril, J. F.; Girones, R.; Bofill-Mas 490 to Human Infectious Diseases. Front Microbiol 2018, 9, 2924.
491 13. Martinez-Puchol, S.; Rusinol, M.; Fernandez-Cassi, X.;
492 Anton, A.; Abril, J. F.; Girones, R.; Bofill-Mas, S., Characterisati
493 NGS tools and occ Anton, A.; Abril, J. F.; Girones, R.; Bofill-Mas, S., Characterisation of the sewage virome: comparisor NGS tools and occurrence of significant pathogens. *Sci Total Environ* **2020**, 713, 136604.
494 14. Crits-Christoph A, 193 NGS tools and occurrence of significant pathogens. *Sci Total Environ* **2020,** 713, 136604.
194 14. Crits-Christoph A, K. R., Olm MR, Whitney ON, Al-Shayeb B, Lou YC, Flamholz A, Kennedy LC,
195 Greenwald H, Hinkle A,

493 NGS tools and occurrence of significant pathogens. Scr *Fotal Environ 2020*, 713, 136604.
494 14. Crits-Christoph A, K. R., Olm MR, Whitney ON, Al-Shayeb B, Lou YC, Flamholz A, K
495 Greenwald H, Hinkle A, Hetzel J, Sp

495 Greenwald H, Hinkle A, Hetzel J, Spitzer S, Koble J, Tan A, Hyde F, Schroth G, Kuersten S, Banfield JF, Greenwald H, Hinkle A, Hetzel J, Spitzer S, Koble J, Tan A, Hyde F, Schroth G, Kuersten S, Banfield JF, 495 Greenwald H, Hinkle A, Hetzel J, Spitzer S, Koble J, Tan A, Hyde F, Schroth G, Kuersten S, Banfield JF,

496 Nelson KL., Genome Sequencing of Sewage Detects Regionally Frevalent SARS-Cov-2 Variants. mblom
197 2021, 12, (1).
198 15. Rothman, J. A.; Loveless, T. B.; Kapcia, J., 3rd; Adams, E. D.; Steele, J. A.; Zimmer-Faust, A. 497 2021, 12, (1).

498 15. Rothn

499 Langlois, K.; W

500 Southern Calif

501 *Microbiol* 202

502 16. Hjelm

503 Elsässer, D.; A

504 of Methods fo 15. Rothman, J. A.; Loveless, T. B.; Kapcia, J., 3rd; Adams, E. D.; Steele, J. A.; Zimmer-Faust, A. G.;

199 Langlois, K.; Wanless, D.; Griffith, M.; Mao, L.; Chokry, J.; Griffith, J. F.; Whiteson, K. L., RNA Viromics of
 500 Southern California Wastewater and Detection of SARS-CoV-2 Single-Nucleotide Variants. Appl Environ
501 Microbiol 2021, 87, (23), e0144821.
502 16. Hjelmsø, M. H.; Hellmér, M.; Fernandez-Cassi, X.; Timoneda, N.; Lukjan Southern California Wastewater and Detection of SARS-CoV-2 Single-Nucleotide Variants. Appr Environt Microbiol 2021, 87, (23), e0144821.

502 16. Hjelmsø, M. H.; Hellmér, M.; Fernandez-Cassi, X.; Timoneda, N.; Lukjancenko, Sol Microbiol 2021, 87, (23), e0144821.
502 16. Hjelmsø, M. H.; Hellmér, M.;
503 Elsässer, D.; Aarestrup, F. M.; Löfstro
504 of Methods for the Concentration ar
505 Sequencing. PLOS ONE 2017, 12, (1),
506 17. Ahmed, W.; Bi Elsässer, D.; Aarestrup, F. M.; Löfström, C.; Bofill-Mas, S.; Abril, J. F.; Girones, R.; Schultz, A. C., Evalua
504 of Methods for the Concentration and Extraction of Viruses from Sewage in the Context of Metagenc
505 Sequ 504 of Methods for the Concentration and Extraction of Viruses from Sewage in the Context of Metagenomic
505 Sequencing. *PLOS ONE* 2017, 12, (1), e0170199.
506 17. Ahmed, W.; Bivins, A.; Metcalfe, S.; Smith, W. J. M.; Ver 505 Sequencing. *PLOS ONE* 2017, 12, (1), e0170199.
506 17. Ahmed, W.; Bivins, A.; Metcalfe, S.; Smith, W. J. M.; Verbyla, M. E.; Symonds, E. M.; Simpson, S.
507 L., Evaluation of process limit of detection and quantificat Sequencing. PLOS ONE 2017, 12, (1), e0170199.
506 17. Ahmed, W.; Bivins, A.; Metcalfe, S.; Smit
507 L., Evaluation of process limit of detection and q
508 dPCR assays for wastewater surveillance. Water
509 18. North, D.; B 19. Ahmed, W.; Statistical and quantification variation of SARS-CoV-2 RT-qPCR and RT-
19. Ahmed, Massays for wastewater surveillance. Water Res 2022, 213, 118132.
18. North, D.; Bibby, K., Comparison of viral concentration 514 a., 160914 . Sociency and CR assays for wastexater surveillance. Water Res 2022, 213, 118132.

509 18. North, D.; Bibby, K., Comparison of viral concentration technique

pathogens from wastewater. Sci Total Environ 2023, 905, 167190.
 510 pathogens from wastewater. *Sci Total Environ* **2023**, *905*, 167190.
511 19. Giron-Guzman, I.; Diaz-Reolid, A.; Cuevas-Ferrando, E.; Falco, I.; Cano-Jimenez, P.; Comas, I.;
512 Perez-Cataluna, A.; Sanchez, G., Evaluat Bathogens from wastewater. Scribture Environ 2023, 909, 107190.
511 19. Giron-Guzman, I.; Diaz-Reolid, A.; Cuevas-Ferrando, E.; Fale
512 Perez-Cataluna, A.; Sanchez, G., Evaluation of two different concer
513 human viruses 19. Giron-Guzman, I.; Diaz-Reolid, A.; Cuevas-Ferrando, E.; Falco, I.; Cano-Jimenez, P.; Comas, I.;

19. Perez-Cataluna, A.; Sanchez, G., Evaluation of two different concentration methods for surveillance of

19. The man v First Transmitter of SARS-CoV-2 sequencing. Sci Total Environ 2023, 862,
514 160914.
515 20. Izquierdo-Lara, R.; Elsinga, G.; Heijnen, L.; Munnink, B. B. O.; Schapendonk, C. M. E.;
516 Nieuwenhuijse, D.; Kon, M.; Lu, L.; A 160914.
513 human viruses in sewage and their effects on SARS-CoV-2 sequencing: Sci Total Environ 2023, 862,
515 20. lzquierdo-Lara, R.; Elsinga, G.; Heijnen, L.; Munnink, B. B. O.; Schapendonk, C. M. E.;
516 Nieuwenhuijse 515 20. |
516 Nieuwen
517 Graaf, M
518 Sequenc
519 21. J
520 Bänziger
521 Jablonsk S16 Nieuwenhuijse, D.; Kon, M.; Lu, L.; Aarestrup, F. M.; Lycett, S.; Medema, G.; Koopmans, M. P

Graaf, M., Monitoring SARS-CoV-2 Circulation and Diversity through Community Wastewater

Sequencing, the Netherlands and Bel Graaf, M., Monitoring SARS-CoV-2 Circulation and Diversity through Community Wastewater

Sequencing, the Netherlands and Belgium. *Emerg Infect Dis* 2021, 27, (5), 1405-1415.

21. Jahn, K.; Dreifuss, D.; Topolsky, I.; Kull 517 Graaf, M., Monitoring SARS-CoV-2 Circulation and Diversity through Community Wastewater Saudion Sequencing, the Netherlands and Belgium. Emerg inject Dis 2021, 27, (3), 1403-1413.

519 21. Jahn, K.; Dreifuss, D.; Topolsky, I.; Kull, A.; Ganesanandamoorthy, P.; Fernande

520 Bänziger, C.; Devaux, A. J.; Stachl Bänziger, C.; Devaux, A. J.; Stachler, E.; Caduff, L.; Cariti, F.; Corzón, A. T.; Fuhrmann, L.; Chen, C.;
521 Jablonski, K. P.; Nadeau, S.; Feldkamp, M.; Beisel, C.; Aquino, C.; Stadler, T.; Ort, C.; Kohn, T.; Julia
522 Be Jablonski, K. P.; Nadeau, S.; Feldkamp, M.; Beisel, C.; Aquino, C.; Stadler, T.; Ort, C.; Kohn, T.; Julia
522 Beerenwinkel, N., Early detection and surveillance of SARS-CoV-2 genomic variants in wastewate
523 COJAC. *Natur* Beerenwinkel, N., Early detection and surveillance of SARS-CoV-2 genomic variants in wastewater using

523 COJAC. *Nature Microbiology* 2022, 7, (8), 1151-1160.

524 22. Jahn, K.; Dreifuss, D.; Topolsky, I.; Kull, A.; Gane 523 COJAC. *Nature Microbiology* 2022, 7, (8), 1151-1160.
524 22. Jahn, K.; Dreifuss, D.; Topolsky, I.; Kull, A.; Ganesanandamoorthy, P.; Fernandez-Cassi, X.;
525 Bänziger, C.; Devaux, A. J.; Stachler, E.; Caduff, L.; Cari S23 COJAC. Nature Microbiology 2022, 7, (0), 1151-1160.
524 22. Jahn, K.; Dreifuss, D.; Topolsky, I.; Kull, A.; Gar
525 Bänziger, C.; Devaux, A. J.; Stachler, E.; Caduff, L.; Cari
526 Jablonski, K. P.; Nadeau, S.; Feldkamp 525 Bänziger, C.; Devaux, A. J.; Stachler, E.; Caduff, L.; Cariti, F.; Corzón, A. T.; Fuhrmann, L.; Chen, C.;
526 Jablonski, K. P.; Nadeau, S.; Feldkamp, M.; Beisel, C.; Aquino, C.; Stadler, T.; Ort, C.; Kohn, T.; Julia
52 Jablonski, K. P.; Nadeau, S.; Feldkamp, M.; Beisel, C.; Aquino, C.; Stadler, T.; Ort, C.; Kohn, T.; Julia
527 Beerenwinkel, N., Detection and surveillance of SARS-CoV-2 genomic variants in wastewater. In C
528 Spring Harbo S27 Beerenwinkel, N., Detection and surveillance of SARS-CoV-2 genomic variants in wastewater. In Cold

Spring Harbor Laboratory: 2021.

23. McCall, C.; Leo Elworth, R. A.; Wylie, K. M.; Wylie, T. N.; Dyson, K.; Doughty, R 528 Spring Harbor Laboratory: 2021.
529 23. McCall, C.; Leo Elworth, R. A.; Wylie, K. M.; Wylie, T. N.; Dyson, K.; Doughty, R.; Treangen, T. J
530 Hopkins, L.; Ensor, K.; Stadler, L. B., Targeted Metagenomic Sequencing for 1982 - Thing Harbor, 1982

529 - 23. McCall, C.; Leo Elworth, F

530 Hopkins, L.; Ensor, K.; Stadler, L. I

532 - 24. Spurbeck, R. R.; Catlin, L.

533 metatranscriptomic methods to Front Public Health 2023, 11, 114

535 - 1922. Hopkins, L.; Ensor, K.; Stadler, L. B., Targeted Metagenomic Sequencing for Detection of Vertebrate

1923. Viruses in Wastewater for Public Health Surveillance. ACS ES&T Water 2023, 3, (9), 2955-2965.

1923. 24. Spur Viruses in Wastewater for Public Health Surveillance. ACS ES&T Water 2023, 3, (9), 2955-2965.

532 24. Spurbeck, R. R.; Catlin, L. A.; Mukherjee, C.; Smith, A. K.; Minard-Smith, A., Analysis of

533 metatranscriptomic meth Viruses in Wastewater for Public Health Surveillance: Acs 25&7 Water 2023, 3, (9), 2555-2565.

24. Spurbeck, R. R.; Catlin, L. A.; Mukherjee, C.; Smith, A. K.; Minard-Smith, A., Analysis of

333 metatranscriptomic methods metatranscriptomic methods to enable wastewater-based biosurveillance of all infectious dise
534 Front Public Health 2023, 11, 1145275.
535 25. Wyler, E.; Lauber, C.; Manukyan, A.; Deter, A.; Quedenau, C.; Teixeira Alves, Front Public Health 2023, 11, 1145275.

535 25. Wyler, E.; Lauber, C.; Manukyan, A.; Deter, A.; Quedenau, C.; Teixeira Alves, L. G.; Seitz, S.;

536 Altmüller, J.; Landthaler, M., Comprehensive profiling of wastewater viro Front Public Health 2023, 11, 1143275.
535 25. Wyler, E.; Lauber, C.; Manukyar
636 Altmüller, J.; Landthaler, M., Comprehe
537 2022.
538 26. d, M. G.; Huggett, J. F., The Digi
639 Publication of Quantitative Digital PCR I
 Altmüller, J.; Landthaler, M., Comprehensive profiling of wastewater viromes by genomic sequencii

537 2022.

26. d, M. G.; Huggett, J. F., The Digital MIQE Guidelines Update: Minimum Information for

7. Publication of Qua 2022.
538 26. d, M. G.; Huggett, J. F., The Digital MIQE Guidelines Update: Minimum Information for
539 Publication of Quantitative Digital PCR Experiments for 2020. *Clin Chem* 2020, 66, (8), 1012-1029.
540 27. B, B., BBT 537 2022.
538 26.
539 Public
540 27.
541 28.
542 Manip 539 Publication of Quantitative Digital PCR Experiments for 2020. *Clin Chem* 2020, 66, (8), 1012-10:
540 27. B, B., BBTools software packag. In *e*, 2014.
541 28. Shen, W.; Le, S.; Li, Y.; Hu, F., SeqKit: A Cross-Platform 1 Concentration of Quantitative Digital PcN Experiments for 2020. Clin Chem 2020, 00, (0), 1012-1029.
540 27. B, B., BBTools software packag. In e, 2014.
541 28. Shen, W.; Le, S.; Li, Y.; Hu, F., SeqKit: A Cross-Platform a 540 27. B, B., BBTools software packag. In *e*, 2014.
541 28. Shen, W.; Le, S.; Li, Y.; Hu, F., SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File
542 Manipulation. *PLOS ONE* **2016,** *11*, (10), e0163962. 542 Manipulation. *PLOS ONE* 2016, 11, (10), e0163962. Manipulation. PLOS ONE 2016, 11, (10), e0163962.

29. Langmead, B.; Salzberg, S. L.; Fast gapped-read alignment with Bowtie 2. Nature Methods 2012,
544 9, (4), 357-359.
545 30. Kim, D.; Song, L.; Breitwieser, F. P.; Salzberg, S. L., Centrifuge: rapid and sensitive classif 944 9, (4), 357-359.

545 30. Kim, D.;

546 metagenomic se

547 31. Martí, J

548 metagenomics.

549 32. Ondov,

550 M., Mash: fast g

546 metagenomic sequences. *Genome Res* 2016, 26, (12), 1721-1729.
547 31. Martí, J. M., Recentrifuge: Robust comparative analysis and contamination removal for
548 metagenomics. *PLOS Computational Biology* 2019, 15, (4), SAD Mictagenomic sequences. Genome Res 2010, 20, (12), 1721-1729.

547 31. Martí, J. M., Recentrifuge: Robust comparative analysis an

548 metagenomics. *PLOS Computational Biology* 2019, 15, (4), e10069

32. Ondov, B. D.; 548 metagenomics. *PLOS Computational Biology* 2019, 15, (4), e1006967.

549 32. Ondov, B. D.; Treangen, T. J.; Melsted, P.; Mallonee, A. B.; Bergman, N. H.; Koren, S.; Ph

550 M., Mash: fast genome and metagenome distance 548 metagenomics. PLOS Computational Biology 2013, 15, (4), e1000507.
549 32. Ondov, B. D.; Treangen, T. J.; Melsted, P.; Mallonee, A. B.; Ber
650 M., Mash: fast genome and metagenome distance estimation using M
651 (1), 1 550 M., Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biology* 2016, 17, (1), 132.
552 33. Oksanen J, S. G., Blanchet F, Kindt R, Legendre P,; Minchin P, O. H. R., Solymos P, Stevens M, Szoecs 551 (1), 132.
551 (1), 132.
552 33. Oksanen J, S. G., Blanchet F, Kindt R, Legendre P,; Minchin P, O. H. R., Solymos P, Stevens M,
553 Szoecs E,; Wagner H, B. M., Bedward M, Bolker B, Borcard D,; Carvalho G, C. M., De Cace

552 33. C

553 Szoecs E,

554 S.; Evang

555 Ribeiro C

556 Package

557 package

-
-

States States Concerts, Wagner H, B. M., Bedward M, Bolker B, Borcard D,; Carvalho G, C. M., De Caceres M, Duran

States, Evangelista H, F. R., Friendly M, Furneaux B,; Hannigan G, H. M., Lahti L, McGlinn D, Ouellette M,;
 554 S.; Evangelista H, F. R., Friendly M, Furneaux B,; Hannigan G, H. M., Lahti L, McGlinn D, Ouellette M,;

855 Ribeiro Cunha E, S. T., Stier A, Ter Braak C, Weedon; Jablonski, K. P., _vegan: Community Ecology

755 Packag 555 Ribeiro Cunha E, S. T., Stier A, Ter Braak C, Weedon; Jablonski, K. P., _vegan: Community Ecology
556 Package_. R
557 package version 2.6-4. 2022.
558 34. Khare, S.; Gurry, C.; Freitas, L.; Schultz, M. B.; Bach, G.; Di Package_. R

S55 package version 2.6-4. 2022.

S58 34. Khare, S.; Gurry, C.; Freitas, L.; Schultz, M. B.; Bach, G.; Diallo, A.; Akite, N.; Ho, J.; Lee, R.

S59 W.; Curation Team, G. C.; Maurer-Stroh, S., GISAID's Role in P 557 package ver

558 34. Khar

559 W.; Curation

560 3, (49), 1049

561 35. Scho

762 McVeigh, R.; 957 Package version 2:0-4. 2022.
558 34. Khare, S.; Gurry, C.; Francess 34. Khare, S.; Gurry, C.; Francess 3, (49), 1049-1051.
561 35. Schoch, C. L.; Ciufo, S.; McVeigh, R.; O'Neill, K.; Robbe
563 Mizrachi, I., NCBI Taxono

W.; Curation Team, G. C.; Maurer-Stroh, S., GISAID's Role in Pandemic Response. *China CDC Wkly* 2021,
560 3, (49), 1049-1051.
561 35. Schoch, C. L.; Ciufo, S.; Domrachev, M.; Hotton, C. L.; Kannan, S.; Khovanskaya, R.; Le W.; Curation Team, G. C.; Maurer-Stroh, S., GISAID's Role in Pandemic Response. *China CDC Wkly* 2021,
560 3, (49), 1049-1051.
561 35. Schoch, C. L.; Ciufo, S.; Domrachev, M.; Hotton, C. L.; Kannan, S.; Khovanskaya, R.; Le 560 5, (49), 1049-1051.
561 35. Schoch, C. L
562 McVeigh, R.; O'Neill
563 Mizrachi, I., NCBI Ta
564 *(Oxford)* 2020, 2020
565 36. Bankevich,
566 Nikolenko, S. I.; Pha
567 M. A.; Pevzner, P. A

561 35. Schoch, C. L.; Ciufo, S.; Domrachev, M.; Hotton, C. L.; Kannan, S.; Khovanskaya, R.; Leipe, D.;
562 McVeigh, R.; O'Neill, K.; Robbertse, B.; Sharma, S.; Soussov, V.; Sullivan, J. P.; Sun, L.; Turner, S.; Karsch-
56 McVeigh, R.; O'Neill, K.; Robbertse, B.; Sharma, S.; Soussov, V.; Sullivan, J. P.; Sun, L.; Turner, S.; Karsch-
563 Mizrachi, I., NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database*
564 *(Oxf* Mizrachi, I., NCBI Taxonomy: a comprenensive apeate on caration, resources and tools. Database

564 (Oxford) 2020, 2020.

565 36. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A. A.; Dvorkin, M.; Kulikov, A. S.; Lesin, V

102012020, 2020.
565 36. Bankevich, A.
566 Nikolenko, S. I.; Phan
567 M. A.; Pevzner, P. A., sequencing. J Compute
569 37. Guo, J.; Boldu
570 A.; Gazitúa, M. C.; Vik
571 to detect diverse DNA Nikolenko, S. I.; Pham, S.; Prjibelski, A. D.; Pyshkin, A. V.; Sirotkin, A. V.; Vyahhi, N.; Tesler, G.; Alekseyev,

567 M. A.; Pevzner, P. A., SPAdes: a new genome assembly algorithm and its applications to single-cell

56 M. A.; Pevzner, P. A., SPAdes: a new genome assembly algorithm and its applications to single-cell

sequencing. J Comput Biol 2012, 19, (5), 455-77.

37. Guo, J.; Bolduc, B.; Zayed, A. A.; Varsani, A.; Dominguez-Huerta, G. 568 sequencing. J Comput Biol 2012, 19, (5), 455-77.

569 37. Guo, J.; Bolduc, B.; Zayed, A. A.; Varsani, A.; Dominguez-Huerta, G.; Delmont, T. O.; Pratam

4.; Gazitúa, M. C.; Vik, D.; Sullivan, M. B.; Roux, S., VirSorter2 569 37. Guo, J.; Bolduc, B.; Zayed, A. A.; Varsani,
569 37. Guo, J.; Bolduc, B.; Zayed, A. A.; Varsani,
570 A.; Gazitúa, M. C.; Vik, D.; Sullivan, M. B.; Roux, S
571 to detect diverse DNA and RNA viruses. *Microbio*
572 38 A.; Gazitúa, M. C.; Vik, D.; Sullivan, M. B.; Roux, S., VirSorter2: a multi-classifier, expert-guided approach
571 to detect diverse DNA and RNA viruses. *Microbiome* 2021, 9, (1), 37.
572 38. Sayers, E. W.; Bolton, E. E.;

to detect diverse DNA and RNA viruses. *Microbiome* 2021, 9, (1), 37.

572 38. Sayers, E. W.; Bolton, E. E.; Brister, J. R.; Canese, K.; Chan, J.; Comeau, D. C.; Connor, R.; Funk, K.;

573 Kelly, C.; Kim, S.; Madej, T.; Ma

571 to detect diverse DNA and RNA viruses. Microbiome 2021, 5, (1), 37.
572 38. Sayers, E. W.; Bolton, E. E.; Brister, J. R.; Canese, K.; Chan, J.; Kelly, C.; Kim, S.; Madej, T.; Marchler-Bauer, A.; Lanczycki, C.; Lathrop
 Kelly, C.; Kim, S.; Madej, T.; Marchler-Bauer, A.; Lanczycki, C.; Lathrop, S.; Lu, Z.; Thibaud-Nissen, F.;
S73 Kelly, C.; Kim, S.; Madej, T.; Marchler-Bauer, A.; Lanczycki, C.; Lathrop, S.; Lu, Z.; Thibaud-Nissen, F.;
S. T Murphy, T.; Phan, L.; Skripchenko, Y.; Tse, T.; Wang, J.; Williams, R.; Trawick, B. W.; Pruitt, K. D.; She
575 S. T., Database resources of the national center for biotechnology information. *Nucleic Acids Res* 20.
576 50, 575 S. T., Database resources of the national center for biotechnology information. Nucleic Acids Res 2022,
576 50, (D1), D20-d26.
577 39. Robinson, J. T.; Thorvaldsdóttir, H.; Winckler, W.; Guttman, M.; Lander, E. S.; Get 575 5. 1., Database resources of the national center for biotechnology information: Mache Acids Res 2022, 576 50, (D1), D20-d26.
577 39. Robinson, J. T.; Thorvaldsdóttir, H.; Winckler, W.; Guttman, M.; Lander, E. S.; Getz, 577 39. Robinson, 577 39. Robinson, 578 J. P., Integrative ge
578 J. P., Integrative ge
579 40. Kearse, M. 580 A.; Markowitz, S.; I
581 integrated and ext
582 data. *Bioinformatic*
583 41. Edgar, R. C

578 J. P., Integrative genomics viewer. *Nat Biotechnol* 2011, 29, (1), 24-6.
579 40. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper,
680 A.; Markowitz, S.; Duran, C.; Thie 578 5.1.1, Integrative genomics viewer. Mat Biotechnol 2011, 29, (1), 24-6.
579 40. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.
580 A.; Markowitz, S.; Duran, C.; Thierer, T.; Ashton, B.; Meintjes, P.; Dr A.; Markowitz, S.; Duran, C.; Thierer, T.; Ashton, B.; Meintjes, P.; Drummond, A., Geneious Basic: an
581 integrated and extendable desktop software platform for the organization and analysis of sequence
582 data. *Bioinfo* 4.; Markowitz, S.; Duran, C.; Thierer, T.; Ashton, B.; Meintjes, P.; Drummond, A., Geneious Basic: an

581 integrated and extendable desktop software platform for the organization and analysis of sequence

582 data. *Bioin* data. Bioinformatics 2012, 28, (12), 1647-9.

582 data. Bioinformatics 2012, 28, (12), 1647-9.

583 41. Edgar, R. C., MUSCLE: multiple sequence alignment with high accuracy and high throughput.

584 *Nucleic Acids Res* 200 382 data. Biomformatics 2012, 28, (12), 1647-9.
583 41. Edgar, R. C., MUSCLE: multiple sequ
584 *Nucleic Acids Res* 2004, 32, (5), 1792-7.
585 42. Castresana, J., Selection of Conserve
786 Phylogenetic Analysis. *Molecular*

Stam *Nucleic Acids Res* 2004, 32, (5), 1792-7.

585 42. Castresana, J., Selection of Conserved Blocks from Multiple Alignments for Their Use in

586 Phylogenetic Analysis. *Molecular Biology and Evolution* 2000, 17, (4), 384 Muchel Acids Res 2004, 32, (3), 1732-7.
585 42. Castresana, J., Selection of Cons
586 Phylogenetic Analysis. *Molecular Biolog*
587 43. Tamura, K.; Stecher, G.; Kumar,
588 11. *Mol Biol Evol* 2021, 38, (7), 3022-30.

586 Phylogenetic Analysis. *Molecular Biology and Evolution* **2000,** 17, (4), 540-552.
587 43. Tamura, K.; Stecher, G.; Kumar, S., MEGA11: Molecular Evolutionary Genetics Analysis \
588 11. *Mol Biol Evol* **2021,** 38, (7), 586 Phylogenetic Analysis. Molecular Biology and Evolution 2000, 17, (4), 540-552.
587 43. Tamura, K.; Stecher, G.; Kumar, S., MEGA11: Molecular Evolutionary Ge
588 11. *Mol Biol Evol* 2021, 38, (7), 3022-3027. 588 11. Mol Biol Evol 2021, 38, (7), 3022-3027. 588 11. Mol Biol Evol 2021, 30, (7), 3022 3027 .

44. Xu, W.; Xu, N.; Zhang, M.; Wang, Y.; Ling, G.; Yuan, Y.; Zhang, P., Nanotraps based on
590 multifunctional materials for trapping and enrichment. Acta Biomater 2022, 138, 57-72.
591 45. Richter, L.; Ksiezarczyk, K.; Pa 591 a. Richter, L.; Ksiezarczyk, K.; Paszkowska, K.; Janczuk-Richter, M.; Niedziolka-Jonsse
592 J.; Los, M.; Holyst, R.; Paczesny, J., Adsorption of bacteriophages on polypropylene labwa
593 reproducibility of phage resear 592 J.; Los, M.; Holyst, R.; Paczesny, J., Adsorption of bacteriophages on polypropylene labware affects the
593 reproducibility of phage research. *Sci Rep* **2021**, 11, (1), 7387.
594 46. Mondal, S.; Feirer, N.; Brockman, reproducibility of phage research. Sci Rep 2021, 11, (1), 7387.

594 46. Mondal, S.; Feirer, N.; Brockman, M.; Preston, M. A.; Teter, S. J.; Ma, D.; Goueli, S. A.; Moorji, S

595 Saul, B.; Cali, J. J., A direct capture met 1593 reproducibility of phage research. Scribep 2021, 11, (1), 7387.

594 46. Mondal, S.; Feirer, N.; Brockman, M.; Preston, M. A.; T

595 Saul, B.; Cali, J. J., A direct capture method for purification and

596 epidemiolo Saul, B.; Cali, J. J., A direct capture method for purification and detection of viral nucleic acid enables

epidemiological surveillance of SARS-CoV-2. Sci Total Environ 2021, 795, 148834.

47. Hatfull, G. F.; Hendrix, R. 596 epidemiological surveillance of SARS-CoV-2. Sci Total Environ 2021, 795, 148834.
597 47. Hatfull, G. F.; Hendrix, R. W., Bacteriophages and their genomes. Curr Opin Virol 2011, 1, (4),
598 298-303.
48. Martinez-Puchol, Epidemiological surveilance of SARS-CoV-2. Sci Potal Environ 2021, 755, 140054.

1976 - 298-303.

1986 - 298-303.

1986 - Martinez-Puchol, S.; Itarte, M.; Rusinol, M.; Fores, E.; Mejias-Molina, C.; A

1986 - Quer, J.; Abri 598 298-303.
599 48. Martinez-Puchol, S.; Itarte, M.; Rusinol, M.; Fores, E.; Mejias-Molina, C.; Andres, C.; Anton, A.
600 Quer, J.; Abril, J. F.; Girones, R.; Bofill-Mas, S., Exploring the diversity of coronavirus in sewa 599 48. M

600 Quer, J.; *A*

601 COVID-19

602 49. K

603 Gross, H.;

604 prevalenc

605 evaluation Quer, J.; Abril, J. F.; Girones, R.; Bofill-Mas, S., Exploring the diversity of coronavirus in sewage during

COVID-19 pandemic: Don't miss the forest for the trees. *Sci Total Environ* 2021, 800, 149562.

49. Khan, M.; Li 601 COVID-19 pandemic: Don't miss the forest for the trees. *Sci Total Environ* 2021, 800, 149562.
602 49. Khan, M.; Li, L.; Haak, L.; Payen, S. H.; Carine, M.; Adhikari, K.; Uppal, T.; Hartley, P. D.; Vasque
603 Gross, H. 601 COVID-19 pandemic: Don't miss the forest for the trees. *Sci Total Environ* **2021,** *800*, 149562.
602 49. Khan, M.; Li, L.; Haak, L.; Payen, S. H.; Carine, M.; Adhikari, K.; Uppal, T.; Hartley, P. D.
603 Gross, H.; Pe 609 52. Mejías-Molina, C.; Pico-Tomàs, A.; Beltran-Rubinat, A.; Martínez-Puchol, S.; Corominas, L.; From The Unity Correction of SARS-CoV-2 variants and other respiratory viruses in the community - A multi-site

evaluation. *One Health* **2023**, 16, 100536.

50. Paul, D.; Kolar, P.; Hall, S. G., A review of the impact of 605 evaluation. One Health 2023, 16, 100536.
606 50. Paul, D.; Kolar, P.; Hall, S. G., A review of the impact of environmental factors on the fate
607 transport of coronaviruses in aqueous environments. *npj Clean Water* 2 606 50. Paul, D.; Kolar, P.; Hall, S. G., A rev
607 transport of coronaviruses in aqueous env
608 51. Pinon, A.; Vialette, M., Survival of
609 52. Mejías-Molina, C.; Pico-Tomàs, A.;
610 Rusiñol, M.; Bofill-Mas, S., Effecti transport of coronaviruses in aqueous environments. *npj Clean Water* **2021,** 4, (1).

608 51. Pinon, A.; Vialette, M., Survival of Viruses in Water. *Intervirology* **2018**, 61, (5), 214-222.

52. Mejías-Molina, C.; Pico-T 608 51. Pinon, A.; Vialette, M., Survival of Viruses in Water. *Intervirology* 2018, 61, 609 51. Mejías-Molina, C.; Pico-Tomàs, A.; Beltran-Rubinat, A.; Martínez-Puchol, S.; Rusiñol, M.; Bofill-Mas, S., Effectiveness of pa 609 51. Pinon, A.; Vialette, M.; Survival of Viruses in Water. Intervirology 2018, 61, (5), 214-222.
609 52. Mejías-Molina, C.; Pico-Tomàs, A.; Beltran-Rubinat, A.; Martínez-Puchol, S.; Corominas, I
610 Rusiñol, M.; Bofill Rusiñol, M.; Bofill-Mas, S., Effectiveness of passive sampling for the detection and genetic
611 characterization of human viruses in wastewater. *Environmental Science: Water Research & Techn*
612 **2023**, 9, (4), 1195-120 611 Characterization of human viruses in wastewater. *Environmental Science: Water Research*
612 **2023,** 9, (4), 1195-1204.
613 S. Strubbia, S.; Schaeffer, J.; Oude Munnink, B. B.; Besnard, A.; Phan, M. V. T.; Nieuw
614 de entry Characterization of human viruses in wastewater. Environmental Science: Wuter Research & Technology

612 **2023,** 9, (4), 1195-1204.

613 S3. Strubbia, S.; Schaeffer, J.; Oude Munnink, B. B.; Besnard, A.; Phan, M. V. 812 2023, 9, (4), 1193-1204.
613 53. Strubbia, S.; Scha
614 de Graaf, M.; Schapendo
615 Metavirome Sequencing
616 *Front Microbiol* 2019, 10
617 54. Levican, J.; Levic
618 surveillance in wastewat
619 55. Rafique, A.; Jian de Graaf, M.; Schapendonk, C. M. E.; Wacrenier, C.; Cotten, M.; Koopmans, M. P. G.; Le Guyader, F. S.,
615 Metavirome Sequencing to Evaluate Norovirus Diversity in Sewage and Related Bioaccumulated Oysters.
616 Front Micro Metavirome Sequencing to Evaluate Norovirus Diversity in Sewage and Related Bioaccumulated Oyster
616 Front Microbiol 2019, 10, 2394.
617 54. Levican, J.; Levican, A.; Ampuero, M.; Gaggero, A., JC polyomavirus circulation Front Microbiol 2019, 10, 2394.

617 54. Levican, J.; Levican, A.; Ampuero, M.; Gaggero, A., JC polyomavirus circulation in one-year

618 surveillance in wastewater in Santiago, Chile. *Infect Genet Evol* 2019, 71, 151-158 617 54. Levican, J.; Levican, A.; k
618 surveillance in wastewater in Sa
619 55. Rafique, A.; Jiang, S. C., (
620 wastewater. *J Water Health* 200
621 56. Oghuan, J.; Chavarria, C
622 Brown, E. L.; Cregeen, S. J.; Deeg
 618 surveillance in wastewater in Santiago, Chile. *Infect Genet Evol* 2019, 71, 151-158.
619 55. Rafique, A.; Jiang, S. C., Genetic diversity of human polyomavirus JCPyV in Southern Califor
620 wastewater. *J Water Health* Surveillance in Wastewater in Santiago, Chile. *Infect Genet Evol* 2013, 71, 151-158.
619 55. Rafique, A.; Jiang, S. C., Genetic diversity of human polyomavirus JCPyV in
620 wastewater. J Water Health 2008, 6, (4), 533-8.
 wastewater. J Water Health 2008, 6, (4), 533-8.

621 56. Oghuan, J.; Chavarria, C.; Vanderwal, S. R.; Gitter, A.; Ojaruega, A. A.; Monserrat, C.; Bauer, C.

622 Brown, E. L.; Cregeen, S. J.; Deegan, J.; Hanson, B. M.; Tisz 621 S6. Oghuan, J.; Chavarria, C.; Vanderwal, S.
622 Brown, E. L.; Cregeen, S. J.; Deegan, J.; Hanson, I
623 W.; Rios, J.; Boerwinkle, E.; Mena, K. D.; Wu, F.,
624 cases in a low-prevalence area. *MedRxiv* 2023.
625 57. Wo France Scanner, J.; Cregeen, S. J.; Deegan, J.; Hanson, B. M.; Tisza, M.; Ocaranza, H. I.; Balliew, J.; Maresso, A.

W.; Rios, J.; Boerwinkle, E.; Mena, K. D.; Wu, F., Wastewater surveillance suggests unreported Mpox

case W.; Rios, J.; Boerwinkle, E.; Mena, K. D.; Wu, F., Wastewater surveillance suggests unreported Mpox

cases in a low-prevalence area. *MedRxiv* 2023.

57. Wolfe, M. K.; Duong, D.; Hughes, B.; Chan-Herur, V.; White, B. J.; B cases in a low-prevalence area. *MedRxiv* 2023.

625 S7. Wolfe, M. K.; Duong, D.; Hughes, B.; Chan-Herur, V.; White, B. J.; Boehm, A. B., Detection of

626 monkeypox viral DNA in a routine wastewater monitoring program. *M* 625 57. Wolfe, M. K.; Duong, D.; Hughes, B.; Ch
626 monkeypox viral DNA in a routine wastewater is
627 58. Rehn, A.; Braun, P.; Knüpfer, M.; Wölfe
628 CoV-2 by sequence hybridization: a comparativ
629 monkeypox viral DNA in a routine wastewater monitoring program. *MedRxiv* 2022.
627 58. Rehn, A.; Braun, P.; Knüpfer, M.; Wölfel, R.; Antwerpen, M. H.; Walter, M. C., Catching SARS-
628 CoV-2 by sequence hybridization: a c 627 S8. Rehn, A.; Braun, P.; Knüpfer, M.; Wölfel, R.; Antwerpen, M. H.; Walter, M. C
628 CoV-2 by sequence hybridization: a comparative analysis. In Cold Spring Harbor Lab
629

628 CoV-2 by sequence hybridization: a comparative analysis. In Cold Spring Harbor Laboratory: 2021.