



UNIVERSITAT^{DE}
BARCELONA

L'exposició a virus presents en l'ambient: descripció del viroma i avaluació del risc

Marta Itarte Sorolla



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- NoComercial – SenseObraDerivada 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - NoComercial – SinObraDerivada 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-NonCommercial-NoDerivs 4.0. Spain License.**

L'exposició a virus presents en l'ambient: descripció del viroma i avaluació del risc

Marta Itarte Sorolla

Tesi doctoral

Abril 2024





UNIVERSITAT DE
BARCELONA

Facultat de Biologia

Departament de Genètica, Microbiologia i Estadística

Secció de Microbiologia, Virologia i Biotecnologia

Programa de Doctorat en Biotecnologia

**L'exposició a virus presents en l'ambient:
descripció del viroma i avaluació del risc**

Memòria presentada per

Marta Itarte Sorolla

Per optar al grau de

Doctora per la Universitat de Barcelona

Doctoranda

Signatura manuscrita de Marta Itarte Sorolla.

Marta Itarte Sorolla

Barcelona, 2024

Directora i tutora

Sílvia Bofill Mas

Directora

Marta Rusiñol Arantegui

Als de casa.

*Si totes haguessin de desaparèixer o de morir
hauria d'aprendre a mirar cap a un cel buit
i veure en l'absoluta foscor quelcom gloriós
tot i que això em costaria una mica més d'un dia o dos...*

Mishima – Qui més estima

Aquesta tesi ha estat finançada amb un ajut predoctoral per a la contractació de personal investigador novell (FI 2020).

Disseny de la portada: Núria Itarte.

Biorender.com s'ha utilitzat per a la creació de les figures d'aquesta tesi.

Agraïments

Des de que vaig començar lo doctorat, no he pogut evitar imaginar-me este moment cada vegada que ha caigut a les meues mans alguna tesi i he acabat xafardejant la part dels agraïments, com tothom fa, no mos enganyem. Encara no me puc creure que haigues arribat lo moment d'escriure estes línies.

I és que qui m'hauria d'haver dit a mi, aquell primer dia que vaig posar un peu al laboratori 8 a mitja entrevista pa fer lo TFG al grup, que 8 anys més tard me trobaria escrivint esta tesi. Però aquí estem, i vull dedicar este trosset a totes les persones que m'han acompanyat durant este viatge. Això sí, me prendré la llibertat de fer-ho en lo meu dialecte, lo faldut, tal com penso i parlo.

Primer de tot, m'agradaria donar infinites gràcies a les meues directores per haver-me brindat l'oportunitat de fer la tesi al grup. Gràcies **Sílvia** per aquella primera entrevista i per haver confiat sempre en mi des del primer moment. Ets una dona referent pa mi, en molts aspectes. Gràcies **Marta** per acceptar ser co-directora d'esta tesi, me fa molta il·lusió ser la teua primera doctoranda. Gràcies pels teus moments TOC de posar ordre, no tant pels teus rampells de tirar coses quet fan segregar dopamines, sinó per posar ordre a la vida. Quina sort tenir-vos a les dos de directores i haver tingut l'oportunitat d'aprendre de vatros, tant científicament com personalment. Gràcies també a **Rosina**, per compartir coneixements i experiència, i per trobar sempre un moment per parlar de ciència.

Me sento molt afortunada d'haver format part d'este grup de donasses a les que admiro enormement. Tant les que acabo de mencionar com les que parlaré a continuació. Gràcies a totes estes dones, me quedo en molts bons moments viscuts durant estos anys, com ara los cafès al sol, les cerveses al bar durant la *golden hour*, les farteres de sushi al meu *buffet* japonès de confiança, los clubs de lectura, les pauses per minjar kikos, lo

xocolate salvador en moments d'emergència, les excursions pa gestionar residus, les colònies vircont, los *escape rooms*... també, gràcies a elles, he après que mai se fan PCRs un divendres, que quan se va de mostreig de vegades les tapes no encaixen en los cubells, que no se pot anar al lab lo dia que toca posar una digital, i que disfrassar-te a carnaval d'un virus que s'està convertint en una amenaça mundial és inapropiat (encara que en aquell moment cap s'imaginava lo que mos esperava en poques setmanes...). Ah! I que los dilluns són menos dilluns si hi ha capítol de crims, o fan gala d'OT, o qualsevol altre programa que mos oferixque bon material pa comentar durant lo moment cafè del dia següent. Però lo més important que he après de vatros és, sobretot, la importància de treballar en equip. I vatros sou lo millor equip que se pot tenir.

Gràcies **Sandra**. Tinc molt marcat lo moment quet vaig conèixer. Después d'haver visitat el lab 8, recordo que vam enganxar-te in fraganti fent una *nested* a la cuina. T'ho he dit moltes vegades, però és la veritat. Has sigut tutora del meu TFG, del meu TFM i de la meua vida en general, i avui estic aquí escrivint això en molt bona part gràcies a tu. Gràcies per tots los consells que m'has donat i per ajudar-me a vore les coses clares quan no les tenia. És tot un goig haver pogut compartir en tu les pipetes de metagenòmica i la passió pels papil·lomavirus, los nostres estimats papil·lomavirus. Gràcies per tot lo que m'has ensenyat durant tot este temps.

Una persona que també m'ha marcat estos anys de tesi és **Eva**. Gràcies pel teu bon humor i pel teu dialecte; almenos quan tu estaves, jo no era l'única persona que parlava en un accent raro. Tot i que ara estigues a l'altra punta de món, aquí continues estant molt present. Sobretot en este últim tram de tesi, no podia parar de sentir-me reflectida en la teua època "pre-tèsica". Lo meme dels *Spider-mans*? Pos això mateix.

No he pogut compartir este final de tesi en Eva, però sí que he tingut lo plaer de viure'l mano a mano en **Cris**. Me farà ràbia no haver pogut comprovar si t'endús el títol oficial a la persona més *short* del departament. Referint-me a estatura eh, intel·lectualment estàs a un nivell superior, i la teua rapidesa en fer bromes ho demostra. Molts ànims en lo quet queda de tesi, encara que jo ja acabo esta etapa, sempre estaré disposta a escoltar-te moltes coses inapropiades.

Una altra persona en qui he pogut compartir les penes del doctorat és **Maria**. Gràcies pels moments de birra i fer que les penes siguen menos penes. Tu encara estàs al principi d'esta aventura, però en quan menos t'ho esperes, estaràs escrivint agraïments com jo ara. I aquí estarem, per recolzar quan calgue! Ja ho saps, jo per tu vaig a València i a on faigue falta.

Gràcies al *Lecturer of the Faculty of Pharmacy*, **Xavi**. També me'n recordo molt de tu pel lab 8 durant aquella primera entrevista i d'haver viscut lo teu final de tesi. Gràcies pels *trapitxeos* de cèl·lules i de virus, per ajudar-me sempre que has pogut i per dir-me venir a fer cultiu aquelles setmanes. En realitat, sé que t'encanta que vinguem a visitar-te.

Gràcies a tots los altres membres que formen o han format part del grup. Gracias **Ayalke**, por estar siempre dispuesto a ayudar con una sonrisa. Gràcies **Aiora**, per haver garantit la supervivència de les cèl·lules i per ser tan *crack* seqüenciant variants de coronavirus. Gràcies als xiquets, **Ignasi** i **Júlia**, per aportar alegria i optimisme sempre a la vida, i per ser tan peculiars (en lo bon sentit de la paraula). Gràcies **Diana** i **Ona**, feu un equip xulíssim i arribareu molt lluny. Gràcies **Marina** i **Lola** pels mil mostrejos compartits. Gràcies **Jorge** i **Marta**, espero tornar a coincidir en vatros en algun moment. Gràcies **Pep** pels infinits *runs* de seqüenciació que hem fet i pels mini atacs de cor cada vegada que sortia alguna actualització nova del programari. Gràcies **Miquel** per tota l'ajuda en les anàlisis de QMRA i per haver intentat fer-mos l'estadística una mica més

fàcil. Gràcies **Natàlia**, la Tecles, per ajudar a analitzar les dades d'amplicons. Gràcies **Rosa** per dedicar-me un somriure sempre que mos creuem pel lab 8.

Gràcies a la resta de membres del departament. Gràcies al nostre membre infiltrat **David**, i als altres entèrics, **Albert**, **Alba** i **Adan**. Gràcies a MARS i a tota la gent de la *Music League*, definitivament heu format part de la banda sonora d'este final de tesi. Gracias **Gloria** por estar siempre dispuesta a tomar una birra, gracias **Julia** por las charlas a horas intempestivas cuando no queda nadie en el departamento, gracias **Dani** por mantenernos siempre informadas de las cosas importantes i gràcies **Raquel** pels teus consells sobre cultiu cel·lular. Gràcies a **Maria**, **Pol**, **Robert**, **Andrés** i **Carlos**. Gràcies a tots pels moments viscuts pels passadissos i zones comunes del departament.

També m'agradaria mencionar a tota aquella gent que ja no està al departament però que han dixat petjada. Gràcies **Edu**, **Gemma**, **Sara M.**, **Geno**, **Manu**, **Ale**, **Sara N.**, **Pedro**, **Carmina**, **Elena** i **Arantxa**.

Sense sortir encara del món de la Facultat, m'agradaria agrair a les **Chubbies** aquells anys de carrera en los que vau ser imprescindibles, com ho continueu sent ara. Gràcies **Berta**, **Elisa**, **Laura**, **Zaro**, **Eva V.**, **Maria**, **Clàudia**, **Carlota** i **Eva M.**. És inevitable passar per la gespeta i no pensar en vatos.

Ara sí, finalment, m'agradaria dedicar estes últimes línies a altres amics i familiars que m'han ajudat a seguir avant des de fora. Gràcies als de l'**auto'12** per donar-me energia sempre que baixo al poble. Gràcies a **Ari**, per ser l'amiga que sempre està, i per tot lo que hem viscut. Te miro i t'admiro. Gràcies **Jordi** per ser lo millor company de pis que se pot tenir (però per detràs d'Ari, claro, los 10 anys de convivència són difícils de superar...).

Gràcies a la meua família de Tarragona. Gràcies **Agustí, Dolors, Irene, Sílvia, iaio i àvia**, per fer-me sentir una més de la família des del primer dia. Molts ànims Irene i Sílvia en lo que vos queda de tesi, ja ho teniu també! Moltes gràcies als **galgos** i a les **galgues**, sou molt bona gent i teniu la millor capital de província de Catalunya.

Gràcies a tots los de casa, a qui va dedicada esta tesi. Gràcies a mons pares, **Paco i Rosa**, per tot lo suport incondicional que m'heu donat i per animar-me sempre a aconseguir tot allò que m'he proposat. Gràcies a man germanes, **Anna i Núria**, per haver sigut lo meu exemple a seguir des de que era menuda. Gràcies a mons cunyats, **Diego i Cristian**, per ser los cunyats més cunyats que se pot tenir. Gràcies a **Júlia, Pol, Blanca i Abril** per fer-me sentir la tia més orgullosa. I gràcies també a ma ueles, la **iaia Maria** i la **iaia Lluïsa**, per ser dones fortes i valentes. En especial, a la iaia Lluïsa, que en menos d'un mes farà 100 anys i, sense dubte, és la dona més forta que conec.

Per últim, gràcies **Ricard**. Gràcies per la paciència infinita, molt probablement ets la persona que ho ha viscut (i patit) de més a prop. Només puc dir que quina sort que s'haiguen trobat los nostres camins. Encara que sigues un traïdor de la ciència, fer esta tesi ha sigut molt més fàcil gràcies a tu. Ara i aquí, t'estimo.

Abril 2024

Continguts

Abreviacions.....	3
Glossari.....	7
Resum de la tesi / Abstract.....	9
Introducció.....	13
1. Virus a l'ambient.....	15
1.1. Vies de transmissió.....	15
1.2. Persistència a l'ambient.....	19
1.3. Exposoma víric, viroma i salut humana.....	20
1.4. Virus emergents i <i>One Health</i>	22
1.5. Vigilància ambiental i epidemiologia basada en aigües residuals.....	24
2. Mètodes per a l'estudi de virus en matrius ambientals.....	27
2.1. Mètodes de mostreig i processament de mostres ambientals.....	27
2.2. Eines tradicionals de detecció de virus en mostres ambientals.....	32
2.3. Noves metodologies: seqüenciació massiva per a la detecció de virus en mostres ambientals.....	35
2.3.1. Seqüenciació massiva amb enriquiment de dianes.....	40
2.3.2. Seqüenciació massiva d'amplicons.....	43
3. Avaluació del risc microbiològic associat a l'exposició a virus.....	46
3.1. Avaluació quantitativa del risc microbiològic.....	48
3.2. Risc ocupacional associat a l'exposició vírica.....	52
Objectius.....	57
Informes dels articles.....	61
1. Llista d'articles inclosos a la tesi.....	63
2. Informe de coautoria.....	65
3. Informe sobre el factor d'impacte dels articles publicats.....	69

Capítol 1. L'aigua residual: font de contaminació i eina per a estudis epidemiològics. Detecció i caracterització de coronavirus en un context de pandèmia.....71

Article 1. Buscant una agulla en un paller. Caracterització de variants de SARS-CoV-2 a les aigües residuals75

Article 2. Explorant la diversitat de coronavirus en les aigües residuals durant la pandèmia de la COVID-19: que els arbres no ens impedeixin veure el bosc.....85

Article 3. Metaviròmica amb enriquiment de dianes per a la vigilància de coronavirus en aigües residuals i mostres d'origen i animal.....95

Capítol 2. Exposició a virus contaminants d'aigües i aliments: detecció i caracterització vírica.....115

Article 4: Les tècniques de NGS revelen una gran diversitat de virus RNA i papil·lomavirus en aliments de producció orgànica i en aigües de reg.....119

Article 5: Explorant la contaminació vírica en aigües subterrànies i d'escorrentia en un entorn urbà.....143

Capítol 3. Identificació de virus presents en entorns de treball amb risc d'exposició. Avaluació del risc ocupacional.....161

Article 6: Avaluació de l'exposició ambiental a virus en estacions depuradores d'aigua residual i granges de porcs mitjançant seqüenciació de nova generació i aproximacions de risc ocupacional.....165

Discussió183

Conclusions.....205

Referències.....209

Annexos241

Material suplementari de l'article 1243

Material suplementari de l'article 2251

Material suplementari de l'article 3259

Material suplementari de l'article 4263

Material suplementari de l'article 5269

Material suplementari de l'article 6277

Altres publicacions293

Abreviacions

ADN	Àcid desoxiribonucleic
ADS	Seqüenciació massiva d'amplicons (de l'anglès <i>amplicon deep sequencing</i>)
ARN	Àcid ribonucleic
ASFV	Virus de la pesta porcina africana (de l'anglès <i>African swine fever virus</i>)
AstV	Astrovirus
CoV	Coronavirus
COVID-19	Malaltia causada pel SARS-CoV-2 (de l'anglès <i>coronavirus disease 2019</i>)
CSFV	Virus de la pesta porcina clàssica (de l'anglès <i>Classical swine fever virus</i>)
DALY	Any de vida ajustat per discapacitat (de l'anglès <i>disability-adjusted life year</i>)
dPCR	PCR digital (de l'anglès <i>digital PCR</i>)
ddPCR	dPCR basada en gotes (de l'anglès <i>droplet digital PCR</i>)
EBV	Virus d'Epstein-Barr (de l'anglès <i>Epstein-Barr virus</i>)
EDAR	Estació depuradora d'aigües residuals
ELISA	Assaig d'immunoabsorció lligat a enzims (de l'anglès <i>enzyme-linked immunosorbent assay</i>)
EMMI	Directrius sobre la informació mínima per reportar dades de qPCR i dPCR en estudis de microbiologia ambiental (de l'anglès <i>Environmental Microbiology Minimum Information</i>)
EPA	Agència de Protecció Ambiental (de l'anglès <i>Environmental Protection Agency</i>)
EPI	Equip de protecció individual

ABREVIACIONS

et al.	I col·laboradors (del llatí <i>et alii</i>)
EV	Enterovirus
GC	Copies genòmiques (de l'anglès <i>genome copies</i>)
HAdV	Adenovirus humà (de l'anglès <i>Human adenovirus</i>)
HAV	Virus de l'hepatitis A (de l'anglès <i>Hepatitis A virus</i>)
HEV	Virus de l'hepatitis E (de l'anglès <i>Hepatitis E virus</i>)
HPV	Papil·lomavirus humà o virus del papil·loma humà (de l'anglès <i>Human papillomavirus</i>)
HTS	Seqüenciació massiva no dirigida o d'alt rendiment (de l'anglès <i>high throughput sequencing</i>)
IAV	Virus de la grip A (de l'anglès <i>Influenza A virus</i>)
ICC-PCR	Cultiu cel·lular integrat a PCR (de l'anglès <i>integrated cell culture-PCR</i>)
LAMP	Amplificació isotèrmica mediada per bucle (de l'anglès <i>loop-mediated isothermal amplification</i>)
MCPyV	Poliomavirus de les cèl·lules de Merkel (de l'anglès <i>Merkel Cell Polyomavirus</i>)
MDA	Amplificació per desplaçament múltiple (de l'anglès <i>Multiple displacement amplification</i>)
MIQE	Directrius sobre la informació mínima per a la publicació d'experiments qPCR (de l'anglès <i>Minimum Information for Publication of Quantitative Real-Time PCR Experiments</i>)
MPXV	Virus de la verola del mico (de l'anglès <i>Monkeypox virus</i>)
NGS	Seqüenciació de nova generació (de l'anglès <i>next-generation sequencing</i>)

NoV	Norovirus
NoV GI	Norovirus del genogrup I
NoV GII	Norovirus del genogrup II
nPCR	PCR niuada (de l'anglès <i>nested PCR</i>)
OMS	Organització Mundial de la Salut
PAdV	Adenovirus porcí (de l'anglès <i>Porcine adenovirus</i>)
PBS	Tampó fosfat salí (de l'anglès <i>phosphate-buffered saline</i>)
PCR	Reacció en cadena de la polimerasa (de l'anglès <i>polymerase chain reaction</i>)
PEDV	Virus de la diarrea epidèmica porcina (de l'anglès <i>Porcine epidemic diarrhea virus</i>)
PRRSV	Virus de la síndrome reproductiva i respiratòria porcina (de l'anglès <i>Porcine reproductive and respiratory syndrome virus</i>)
QMRA	Avaluació quantitativa del risc microbiològic (de l'anglès <i>quantitative microbial risk assessment</i>)
qPCR	PCR quantitativa (de l'anglès <i>quantitative PCR</i>)
RoV	Rotavirus
RT	Retrotranscripció
SARS-CoV-2	Coronavirus 2 de la síndrome respiratòria aguda greu (de l'anglès <i>Severe acute respiratory syndrome-related coronavirus 2</i>)
SCC	Carcinoma de cèl·lules escamoses (de l'anglès <i>squamous cell carcinoma</i>)
SISPA	Amplificació independent de seqüència basada en un únic encebador (de l'anglès <i>sequence-independent single-primer amplification</i>)
TES	Seqüenciació massiva amb enriquiment de dianes (de l'anglès <i>target enrichment sequencing</i>)

ABREVIACIONS

UV	Llum ultraviolada
VIH	Virus de la immunodeficiència humana
VOC	Variant preocupant del SARS-CoV-2 (de l'anglès <i>variant of concern</i>)
VOI	Variant d'interès del SARS-CoV-2 (de l'anglès <i>variant of interest</i>)
WBE	Epidemiologia basada en aigües residuals (de l'anglès <i>wastewater-based epidemiology</i>)
WWTP	Estació depuradora d'aigües residuals (de l'anglès <i>wastewater treatment plant</i>)

Glossari

Assemblatge	Alineament i fusió de <i>reads</i> per tal de construir <i>contigs</i> .
Avaluació del risc	Caracterització qualitativa o quantitativa i estimació dels possibles efectes adversos per a la salut associats amb l'exposició d'individus o poblacions a un perill.
Bioaerosols	Partícules amb una grandària inferior a 100 µm en suspensió a l'aire que contenen microorganismes.
Comunicació del risc	Intercanvi interactiu d'informació sobre l'anàlisi de risc entre les parts interessades, considerant la percepció pública i la capacitat d'intercanvi d'informació científica.
Contig	Fragment d'ADN obtingut a partir de la superposició d'un conjunt de <i>reads</i> .
Coverage	Cobertura del genoma, percentatge del genoma complet recuperat com a resultat de la seqüenciació massiva.
Exposoma víric	La totalitat de les exposicions a virus que té un individu al llarg de la seva vida, i com aquestes exposicions es relacionen amb el benestar.
Fòmit	Objecte inert que pot esdevenir un focus de transmissió d'infeccions com a conseqüència dels microorganismes patògens que té adherits a la superfície.
Gestió del risc	Procés per controlar un risc determinat, valorant alternatives i seleccionant accions adequades, tenint en compte l'avaluació del risc, així com aspectes d'enginyeria i qüestions econòmiques, legals i polítiques.
Llibreria (seqüenciació)	Conjunt de fragments d'àcids nucleics que, després d'haver passat per una sèrie de passos (unió d'índexs, lligació d'adaptadors, etc.) estan llestos per ser seqüenciats.

Metagenòmica	Terme utilitzat per primera vegada per Jo Handelsman et al. per descriure “l’anàlisi funcional, i basat en seqüències, dels genomes microbians col·lectius que es troben en una mostra ambiental” (Handelsman et al., 1998; Riesenfeld et al., 2004).
Microbioma	Comunitat de microorganismes que existeixen en un entorn particular.
One Health	Enfocament integrador que té com a objectiu equilibrar i optimitzar de manera sostenible la salut dels animals, l’ambient i les persones.
Read	Seqüència obtinguda a partir d’un procés de seqüenciació.
Reassortiment	Procés de formació d’un virus reassortit, agrupant i ordenant segments gènics de virus diferents que han infectat alhora un mateix hoste.
Run (seqüenciació)	Execució del procés de seqüenciació.
Seqüenciació	Deducció de l’ordre dels nucleòtids en una molècula de DNA o RNA per una sèrie de reaccions químiques.
Viroma	Component víric que forma part del microbioma.
Virus emergent	Virus identificat recentment, que està augmentant la seva incidència i/o àmbit geogràfic, o se n’ha descrit noves vies de transmissió.
Virus reassortit	Virus de genoma híbrid resultat de l’agrupament i l’ordenació de segments gènics d’altres virus de soques diferents que han infectat alhora un mateix hoste.
Virus zoonòtic	Virus que es transmet de forma natural d’un animal a un humà.
Zoonosi	Malaltia infecciosa que pot ser transmesa d’un animal a un humà.

Resum de la tesi / Abstract

Existeix una gran diversitat de virus presents en l'ambient als quals els humans estem exposats, constituint l'exposoma víric. Les eines de vigilància ambiental, com ara l'epidemiologia basada en aigües residuals, han posat de manifest el seu potencial per monitoritzar la circulació vírica en una comunitat, sobretot amb la recent pandèmia de la COVID-19. Per aquest motiu, aquesta tesi s'ha centrat, primerament, en l'estudi de l'aigua residual com a font de contaminació i eina per a estudis epidemiològics. Concretament, s'han avaluat diferents mètodes de detecció i caracterització dels coronavirus presents en les aigües residuals en un context de pandèmia, principalment mitjançant tècniques de seqüenciació de nova generació (NGS).

Els virus excretats a les aigües residuals poden ser resistents als tractaments de les estacions depuradores d'aigües residuals (EDARs) i acabar contaminant les fonts d'aigua que s'utilitzen per beure, per regar o per a ús recreatiu, suposant així potencials vies de transmissió d'aquests virus. En aquesta tesi, s'ha estudiat l'exposició a virus patògens presents en aigües i aliments mitjançant la detecció i caracterització vírica, especialment aplicant tècniques de NGS en aliments de producció orgànica i les seves fonts d'aigua de reg. A més, també s'ha avaluat la contaminació vírica i s'ha explorat la diversitat de virus presents en esorrenties i aigües subterrànies d'un entorn urbà, les quals, en l'actual context de sequera, són de gran interès per a diferents usos.

Finalment, considerant que l'exposició ocupacional a patògens pot comportar riscos per a la salut dels treballadors, s'ha avaluat l'exposició vírica en entorns laborals. S'ha quantificat la contaminació vírica i s'han identificat virus presents en aerosols i superfícies d'una EDAR i una granja de porcs mitjançant tècniques de NGS. A més, s'han realitzat estimacions

del risc ocupacional associat a l'exposició a virus amb l'avaluació quantitativa del risc microbiològic (QMRA), seleccionant l'adenovirus humà (HAdV) per a l'EDAR i un virus hipotèticament zoonòtic amb característiques similars a l'adenovirus porcí (PAdV) per a la granja de porcs com a patògens de referència. Les vies d'exposició considerades en els models són la inhalació d'aerosols i la ingestió oral a través de superfícies contaminades i el contacte mà-boca.

Aquesta tesi ha permès la identificació de virus presents en diverses matrius ambientals, incloent-hi aigües residuals, aigües d'escorrentia, aigües subterrànies, mostres d'origen animal, aliments de producció orgànica i aigües de reg, així com superfícies i aerosols d'entorns de treball. S'ha avaluat el nivell de contaminació vírica mitjançant la quantificació de virus indicadors de contaminació fecal i altres virus patògens. S'han identificat virus humans i animals d'interès aplicant tècniques de NGS, com la seqüenciació massiva d'amplicons i la seqüenciació amb enriquiment de dianes, utilitzant panells específics dirigits a la captura de seqüències de virus de vertebrats o també de coronavirus. Finalment, els resultats de QMRA han revelat un potencial risc ocupacional associat a l'exposició a HAdV per als treballadors d'EDARs, si no es prenen mesures. La simulació de QMRA a la granja de porcs també suggereix un potencial risc en cas d'exposició a un virus potencialment zoonòtic.

Els resultats d'aquesta tesi ofereixen una visió rellevant sobre la presència de virus en diverses matrius ambientals i els potencials riscos associats amb l'exposició a aquests, contribuint al desenvolupament de mesures que ajudin a reduir aquests riscos i destacant la importància del monitoratge i la vigilància ambiental per a la detecció de virus emergents i zoonòtics, així com per a entendre les seves rutes de transmissió i identificar possibles fonts de contaminació.

Abstract

There is a wide diversity of viruses present in the environment, constituting the viral exposome to which humans are exposed. Environmental surveillance tools, such as wastewater-based epidemiology, have demonstrated their potential in monitoring viral circulation in a community, particularly with the recent COVID-19 pandemic. For this reason, this thesis has primarily focused on studying wastewater as a source of contamination and as a tool for epidemiological studies. Specifically, different methods for detecting and characterizing coronaviruses in wastewater have been evaluated in the context of a pandemic, particularly through next-generation sequencing (NGS) techniques.

Viruses excreted in wastewater can be resistant to treatments in wastewater treatment plants (WWTPs) and contaminate water sources used for drinking, irrigation, or recreational purposes, thereby becoming potential transmission routes for these viruses. This thesis has aimed to study exposure to pathogenic viruses present in water and food through viral detection and characterization, especially by applying NGS techniques in fresh produce and their irrigation water sources. Additionally, the viral diversity in urban groundwater and runoff has been explored, given its potential uses in the current context of water scarcity.

Furthermore, considering that occupational exposure to pathogens can pose health risks, viral exposure in workplace scenarios has been assessed. Viral contamination has been evaluated, and NGS techniques have been performed to identify the viruses present in aerosols and surfaces to which workers in a WWTP and a swine farm might be exposed. Quantitative Microbial Risk Assessment (QMRA) has been conducted to estimate the occupational risk associated with viral exposure for WWTP workers, choosing human adenovirus (HAdV) as the reference pathogen. In the swine farm, QMRA has been performed as an extrapolation,

considering a hypothetical zoonotic virus with characteristics similar to porcine adenovirus (PAdV). The modelled exposure routes include aerosol inhalation and oral ingestion through contaminated surfaces and hand-to-mouth contact.

This thesis has allowed the identification of viruses in various environmental matrices, including wastewater, runoff, groundwater, animal origin samples, fresh produce, irrigation water, as well as surfaces and aerosols in workplace scenarios. Viral contamination has been evaluated through the quantification of indicators of fecal contamination and other pathogenic viruses. Human and animal viruses of interest have been identified using NGS techniques, such as amplicon deep sequencing and target enrichment sequencing, using specific panels designed to capture sequences of vertebrate viruses or coronaviruses. Finally, QMRA analyses have revealed noteworthy risks associated with exposure to HAdV for WWTP workers if measures are not implemented. QMRA simulation at the swine farm suggests a potential risk in the case of exposure to a hypothetical zoonotic virus.

Overall, the results of this thesis provide relevant insights into the presence of viruses in environmental matrices and the potential risks associated with the exposure to them. This contributes to the development of measures to mitigate these risks and emphasizes the importance of environmental monitoring and surveillance for the detection of emerging and zoonotic viruses, as well as for understanding their transmission routes and identifying potential sources of contamination.

Introducció



1. Virus a l'ambient

Els virus són les entitats més abundants del nostre planeta. S'estima que hi ha 10^{31} virus a la Terra, la gran majoria dels quals són bacteriòfags (Breitbart & Rohwer, 2005). La partícula vírica està constituïda per un genoma, ja sigui àcid ribonucleic (ARN) o àcid desoxiribonucleic (ADN), de simple o doble cadena, portador de la informació genètica, i una càpsida proteica que assegura la seva protecció a l'ambient. En alguns casos, la càpsida pot estar envoltada per una segona capa externa, l'embolcall, format generalment per lípids i glicoproteïnes (Bouseettine et al., 2020). A més de la seva abundància, els virus destaquen per la seva enorme diversitat en l'organització del seu material genètic, els mecanismes de replicació i les interaccions amb els seus hostes, ja siguin antagòniques, comensals o mutualistes (Roossinck, 2015).

1.1. Vies de transmissió

Els virus que s'excreten a l'ambient per un individu infectat poden infectar nous hostes (Figura 1). Per exemple, els virus respiratoris es transmeten mitjançant excrecions respiratòries, és a dir, a través de les gotes o partícules aerosolitzades que es generen quan una persona esternuda, tus, parla o respira (Louten, 2016). Una simple estossada pot generar fins a 10,000 petites partícules d'aerosols potencialment infeccioses, i un esternut pot produir prop de 2 milions (Chappell & Dermody, 2015). La distribució d'aquestes partícules depèn de diversos factors, els quals es detallen més endavant (Figura 2) i, en alguns casos, aquests virus són capaços de viatjar llargues distàncies a través de l'aire (Labadie et al., 2020). Alguns dels virus transmesos per l'aire més representatius són el virus de la grip, el SARS-CoV-2, el virus de la xarampió o el virus respiratori sincicial (Wang et al., 2021).

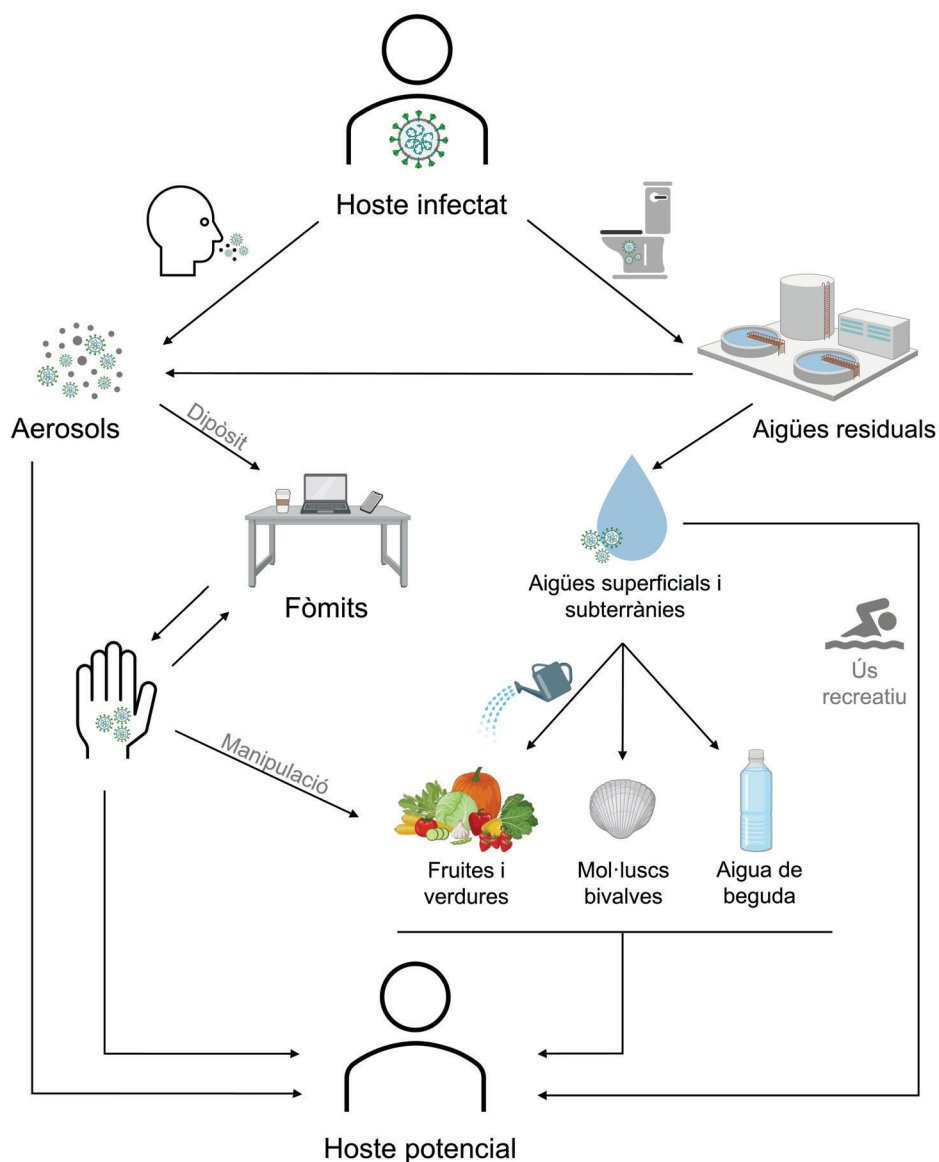


Figura 1. Principals vies de transmissió de virus excretats a l'ambient.

En canvi, els virus que infecten les cèl·lules gastrointestinals s'excreten en femta i segueixen la ruta fecal-oral. Els aliments, l'aigua o les mans contaminades amb matèria fecal poden ser una via de transmissió d'aquests virus, generalment sense embolcall, la qual cosa els confereix una gran estabilitat a l'ambient i, quan arriben al tracte gastrointestinal, els

permet sobreviure a l'acidesa de l'estomac, iniciant així el seu cicle infecciós. Aquests virus suposen problemes importants per a la salut pública, ja que són responsables de malalties com ara gastroenteritis, hepatitis i meningitis (Bouseettine et al., 2020).

Com que els virus de transmissió fecal-oral són excretats en grans quantitats a través de la femta, fins a 10^{11} partícules víriques per gram de femta, aquests es troben en gran abundància a les aigües residuals. El grau de contaminació vírica de les aigües residuals depèn de l'estació de l'any, la prevalença de les infeccions víriques i les característiques dels virus circulants. Els sistemes de tractament d'aigües residuals, fins i tot quan funcionen adequadament, només eliminen aproximadament entre el 20 i el 80% dels virus de transmissió fecal-oral, degut a que aquests són molt resistents als tractaments de desinfecció i poden persistir a l'ambient durant molt temps sense ser desactivats (Biziagos et al., 1988). Aquest fet implica una potencial font de contaminació cap a l'ambient, incloent-hi rius, aigües subterrànies i mars (La Rosa et al., 2012), i aquesta disseminació es pot veure incrementada en períodes de pluja que causin desbordaments de clavegueram i esorrenties (Rusiñol, 2023; Rusiñol et al., 2015). A més, alguns processos que es duen a terme durant el tractament de les aigües residuals poden causar l'aerosolització de les partícules víriques (Corpuz et al., 2020; Kataki et al., 2022).

Així doncs, els virus excretats en femta, però també en orina, saliva i descamacions epitelials, poden acabar contaminant les fonts d'aigua que s'utilitzen per beure, per regar o per a ús recreatiu, constituint d'aquesta manera vies de transmissió d'aquests virus. Els principals virus d'interès associats a malalties transmeses per l'aigua són els norovirus (NoV), els adenovirus humans (HAdV), els astrovirus (AstV), els enterovirus (EV), els rotavirus (RoV), el virus de l'hepatitis A (HAV) i el virus de l'hepatitis E (HEV) (Gibson, 2014).

Els virus de transmissió fecal-oral, a més del consum d'aigua contaminada, també es poden transmetre a través dels aliments. Una gran quantitat de malalties d'origen alimentari s'identifiquen sovint amb etiologia vírica, sent els NoV, HAV, HEV, RoV, AstV i HAdV els virus patògens alimentaris més importants (Koopmans & Duizer, 2004; Newell et al., 2010; Pachepsky et al., 2011; Shaheen et al., 2019). Els aliments poden contaminar-se mitjançant l'aigua de reg, si es tracta de fruites o verdures, o també per la manipulació durant el seu processament (cultiu, collita, emmagatzematge, transport, venda, preparació, etc.) (Koopmans & Duizer, 2004).

D'entre els aliments amb probabilitat de ser contaminats per virus també hi ha els mol·luscs bivalves, com ara els musclos i les ostres. Tot i que aquesta contaminació també pot estar associada a la manipulació, la font principal de contaminació seria, en aquests cas, la contaminació del mar amb aigua residual, ja sigui per fuga de clavegueres i esorrenties durant períodes de pluja o també per aigües residuals tractades de manera deficient (Koopmans & Duizer, 2004; Landrigan et al., 2020). Els mol·luscs bivalves es nodreixen filtrant aigua, acumulant patògens continguts a l'aigua, com són els virus de transmissió fecal-oral. Degut a que sovint es consumeixen crus o poc cuits, esdevenen vehicles eficients per a la transmissió d'aquests virus (Lees, 2000).

A més de la transmissió de virus a través de l'aire, l'aigua i els aliments, també pot produir-se a través de superfícies o fòmits, objectes inerts que poden esdevenir una via de transmissió tant de virus respiratoris com de virus de transmissió fecal-oral. Aquests fòmits es poden contaminar per la deposició de partícules d'aerosols que contenen els virus, així com pel contacte directe amb mans o altres objectes contaminats. D'aquesta manera, un individu que toca aquests fòmits contaminats i posteriorment es toca la cara i/o alguna membrana mucosa pot ser susceptible a ser infectat (Zambrana & Boehm, 2023).

1.2. Persistència a l'ambient

Els virus es caracteritzen per ser paràsits intracel·lulars obligats que probablement infecten totes les formes de vida (Simmonds et al., 2017) i, com hem vist, l'ambient té un paper important com a intermediari o reservori en la seva transmissió. El seu estudi i importants descobriments, com ara la persistència del virus de la grip a l'aire en funció de la mida de les gotes que s'expulsen en tossir o esternudar, o el paper de l'aigua en la transmissió del poliovirus, han contribuït de forma significativa a la millora de la salut pública (Labadie et al., 2020). De la mateixa manera, la recent pandèmia de la COVID-19 també ha demostrat la necessitat d'estudiar la persistència dels virus en el nostre entorn.

L'eficàcia de la transmissió vírica a través de l'ambient depèn de la capacitat de supervivència dels virus en aquest ambient. Diversos factors, sovint interconnectats, determinen la seva persistència, que pot durar des d'uns segons fins a diversos anys. La temperatura, la humitat, el pH, la salinitat de l'aigua o la presència de llum ultraviolada (UV) són factors ambientals que influeixen en la supervivència dels virus (Figura 2). A més, factors vírics com la presència de mutacions en les proteïnes estructurals o la composició dels lípids dels virus embolcallats també hi juguen un paper determinant (Louten, 2016). La virologia ambiental i l'estudi d'aquests factors que afecten la persistència dels virus han contribuït de manera significativa a la millora de la salut de la població, mitjançant el desenvolupament de mètodes de desinfecció, els tractaments d'aigua residual, pràctiques generals d'higiene i la vigilància ambiental (Labadie et al., 2020).

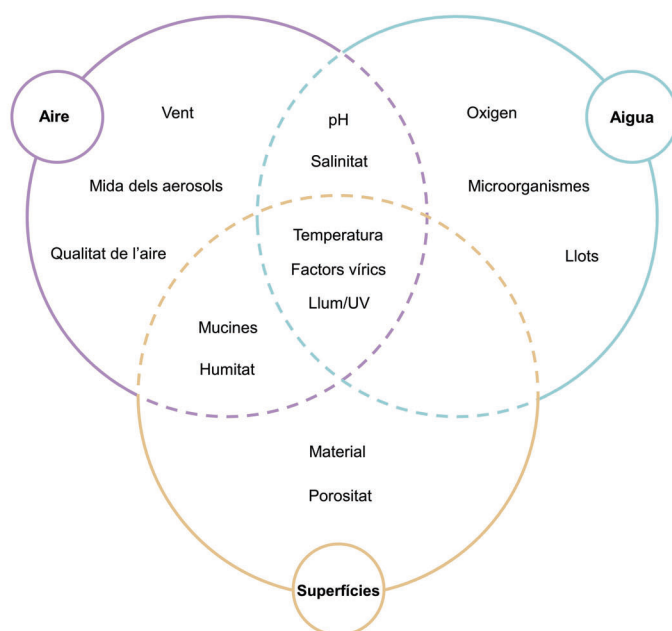


Figura 2. La persistència dels virus a l'ambient depèn de múltiples factors que sovint estan interconnectats. Adaptada de Labadie et al., 2020.

1.3. Exposoma víric, viroma i salut humana

El terme exposoma va ser utilitzat per primera vegada l'any 2005 dins del camp de l'epidemiologia, i es pot definir com “la totalitat de les exposicions que té un individu al llarg de la seva vida, i com aquestes exposicions es relacionen amb el benestar” (Grafanaki et al., 2023; Wild, 2005). Dins d'aquest terme s'inclou l'exposició ambiental a virus, coneguda com a exposoma víric (Gordon-Lipkin et al., 2022).

Al mateix temps, el cos humà comprèn un microecosistema format per una àmplia comunitat de microorganismes, com ara virus, bacteris, arquees i fongs. El component víric que forma part d'aquest microbioma és el viroma humà (Lecuit & Eloit, 2013). Tot i que la majoria dels virus humans millor caracteritzats causen infeccions agudes i estan associats a malalties, un individu sa és hoste d'una varietat molt més gran de virus.

La major part d'aquests infecten bacteris que habiten a l'intestí humà, però alguns virus poden reproduir-se en cèl·lules humanes sense causar malalties, almenys a curt termini i en individus sans. El viroma pot contribuir significativament a l'estat de salut del seu hoste, tot i que sovint és difícil assegurar que cap d'aquests virus no pugui causar, a la llarga, una determinada malaltia sota cap condició. Un exemple clar són els herpesvirus, els quals són pràcticament ubics en la població humana, mantenint-se com a simbiotes en la majoria de les persones la major part del temps, però també poden causar malalties que, en alguns casos, poden ser devastadores en persones immunodeprimides (Koonin et al., 2021).

Existeix una interacció complexa entre l'ambient i aquest viroma humà endogen, la qual està mediada pel sistema immunològic i pot tenir un profund impacte en la salut humana (Figura 3). Entendre aquesta relació és important per conèixer de quina manera influeix en la salut d'un individu i, així, poder desenvolupar noves eines terapèutiques (Sharma & Gilbert, 2018).

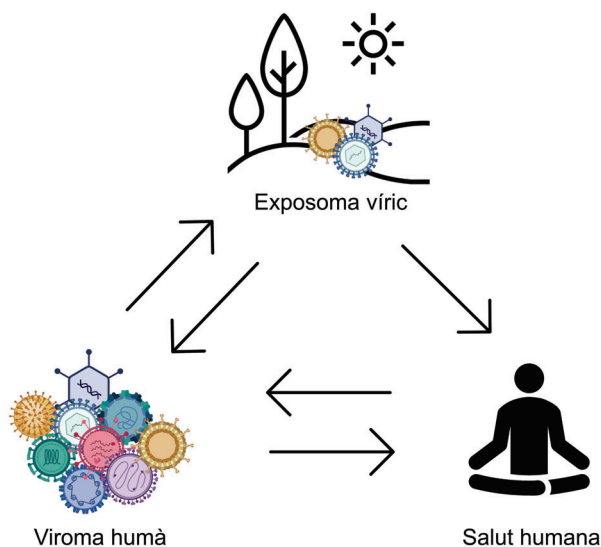


Figura 3. Exposoma víric, viroma humà i salut humana.

1.4. Virus emergents i *One Health*

Els virus emergents són definits com “aquells virus que s’han identificat recentment en una població o que ja existien però que estan augmentant ràpidament la seva incidència i/o àmbit geogràfic, o aquells per als quals s’han descrit noves vies de transmissió” (EPA, 2023; La Rosa et al., 2012). Les relacions complexes entre el virus patògen, l’hoste, i l’ambient determinen l’aparició de virus patògens emergents. En un món globalitzat, els agents infecciosos tenen la capacitat de propagar-se ràpidament per qualsevol part del món. D’entre ells, els virus són els més aptes per esdevenir patògens emergents, ja que poden adaptar-se no només per mutacions, sinó també a través de la recombinació i el reassortiment, adquirint així la capacitat d’infectar nous hostes i adaptar-se a nous entorns (La Rosa et al., 2012).

L’inici del segle XXI s’ha vist marcat per un augment en el nombre de malalties infeccioses emergents detectades arreu del món, la majoria causades per virus zoonòtics, d’origen animal, amb una llarga història d’adaptació als seus hostes naturals i amb capacitat de travessar barreres entre espècies, i amb un immens potencial de disseminació (A. Ahmed et al., 2022; Ermonval & Morand, 2023). El virus de la immunodeficiència humana (VIH), el virus de l’Ebola, el SARS-CoV-2 i el virus de la verola del mico (MPXV) són exemples de virus zoonòtics que s’han associat amb una transmissió inicial d’animals a humans (Bardhan et al., 2023; Louten, 2016).

El SARS-CoV-2 ha estat el causant de la recent pandèmia de la COVID-19, representant una amenaça per a la salut global a causa de les seves altes taxes de transmissió i la pèrdua d’un elevat nombre de vides. A més, també s’ha posat de manifest el potencial del virus per alterar ràpidament l’economia global (Han et al., 2023). Tot i que al maig de 2023 es va donar

la crisi sanitària per finalitzada, el virus segueix circulant i moltes persones continuen infectant-se actualment a Catalunya (Gencat, 2024).

No obstant, la COVID-19 no ha estat l'única malaltia emergent que ha representat una amenaça per a la salut pública en els últims anys. Des del maig de 2022, s'han produït múltiples brots de MPXV a tot el món. La verola del mico va ser declarada una emergència de salut pública per l'Organització Mundial de la Salut (OMS), ja que es feia cada vegada més transmissible entre humans (Han et al., 2023), fins que l'emergència es va donar per acabada al maig de 2023. A Catalunya, en data de 20 de novembre de 2023, s'havien notificat un total de 3,256 casos de verola del mico des dels primers brots (Gencat, 2023).

La recent pandèmia de la COVID-19 i els últims brots de MPXV han posat de manifest la vulnerabilitat de la població davant dels virus zoonòtics. La propagació d'aquests virus està associada a factors ambientals, el canvi climàtic, la salut animal, així com altres activitats humanes, incloent-hi la globalització, la urbanització i els viatges (Shaheen, 2022). És evident que vivim en un món amb un equilibri delicat entre animals, el medi ambient i els éssers humans (Figura 4), la qual cosa reflecteix la necessitat d'una aproximació *One Health* (Una Sola Salut) per al control i la prevenció de malalties zoonòtiques (Sheek-Hussein et al., 2023). Aquest concepte integrador té com a objectiu equilibrar i optimitzar de manera sostenible la salut animal, ambiental i humana, involucrant diverses disciplines i tota la societat per respondre a amenaces per a la salut i el medi ambient (Shaheen, 2022).

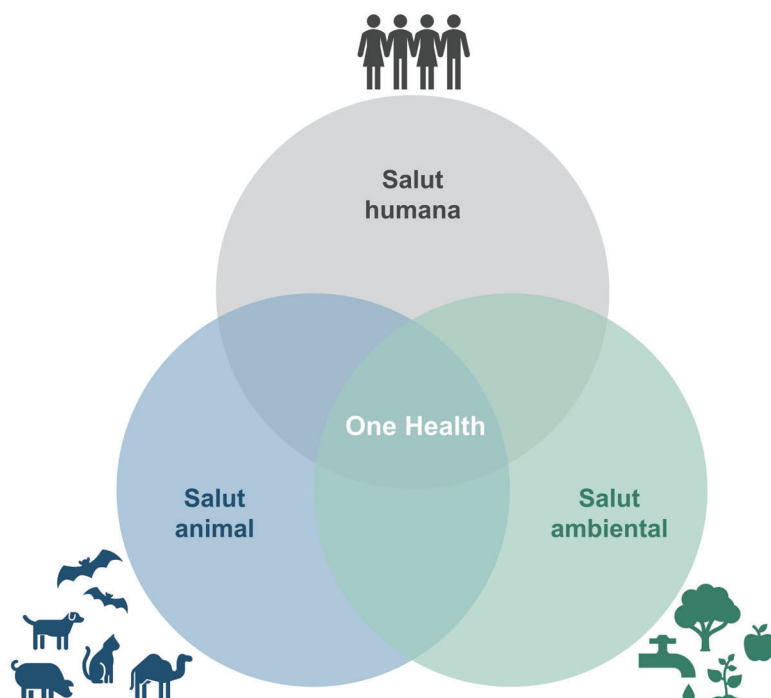


Figura 4. *One Health integra la salut animal, ambiental i humana.*
Adaptada de Ulrich et al., 2023.

1.5. Vigilància ambiental i epidemiologia basada en aigües residuals

Un dels reptes més grans durant la pandèmia de la COVID-19 ha estat el desenvolupament d'eines de vigilància ambiental del SARS-CoV-2, com ara l'epidemiologia basada en aigües residuals (Abdeldayem et al., 2022; Han et al., 2023; Mackul'ak et al., 2021). Aquesta eina ha demostrat el seu potencial com a mètode complementari per avaluar l'evolució de la pandèmia en casos en què la capacitat d'efectuar proves clíniques individuals a la població és limitada, sobretot en aquelles àrees on els recursos són escassos (Hu et al., 2021; Robotto et al., 2022).

El monitoratge de la pandèmia a través de les aigües residuals ha estat possible degut a que el virus s'excreta en femta i el seu ARN és detectable en aquestes, tot i que no s'ha confirmat la presència de partícules víriques infeccioses (Abdeldayem et al., 2022). Molts estudis han demostrat una relació entre les concentracions d'ARN del SARS-CoV-2 detectades a les aigües residuals i el nombre de casos de COVID-19 notificats a la comunitat, tot indicant que la vigilància de les aigües residuals és una font útil d'informació que permet monitoritzar la circulació del virus en la població (Ahmed et al., 2020a; Medema et al., 2020; Rusiñol et al., 2021; Shah et al., 2022). A Catalunya, aquesta eina ha permès la vigilància del SARS-CoV-2 i les seves variants en les aigües residuals, contribuint així a conèixer l'estat i l'evolució de la pandèmia, proporcionant informació per a la detecció anticipada de nous rebrots o l'entrada i circulació de noves variants al territori (ASPCAT, 2022).

Abans de la pandèmia de la COVID-19, l'epidemiologia basada en aigües residuals s'havia utilitzat per monitoritzar la presència de residus farmacèutics o industrials, el consum de drogues en una comunitat i la prevalença d'altres virus (Robotto et al., 2022; Singer et al., 2023). La seva implementació va ser important per al monitoratge del poliovirus com a estratègia mundial per a l'erradicació de la poliomielitis, i encara ara segueix en actiu de forma rutinària en països amb risc d'emergència o re-emergència (Kitakawa et al., 2023; Shah et al., 2022). Alguns exemples d'altres virus patògens que també s'han detectat mitjançant estudis d'epidemiologia d'aigües residuals són el virus de la grip (Wolfe et al., 2022), HAV i NoV (Hellmér et al., 2014), HEV (Iaconelli et al., 2020) i MPXV (de Jonge et al., 2022; Tiwari et al., 2023). A més dels estudis de detecció de patògens que circulen a la població en moments determinants, l'anàlisi de virus presents en les aigües residuals proporciona informació rellevant sobre els virus que infecten la població

de manera persistent i que poden actuar d'indicadors de contaminació fecal (Bofill-Mas et al., 2000; Pina et al., 1998).

Tot i que l'epidemiologia basada en aigües residuals s'ha utilitzat àmpliament per a l'estudi de virus, altres compartiments ambientals involucrats amb la transmissió vírica han rebut menys atenció. És important comprendre les vies de transmissió dels virus per identificar dianes de vigilància ambiental, com ara l'aigua, l'aire, el sòl o les superfícies (Anand et al., 2021; Cela-Dablanca et al., 2021; Núñez-Delgado, 2020). El monitoratge de virus en altres matrius ambientals diferents a les aigües residuals requereix de tecnologies avançades i més estudis per a la seva extensió i aplicabilitat (Abdeldayem et al., 2022).

2. Mètodes per a l'estudi de virus en matrius ambientals

L'estudi dels virus en mostres ambientals és important per identificar les vies de transmissió, detectar brots, caracteritzar la contaminació en una gran varietat d'entorns i estudiar el risc que suposa l'exposició a aquesta. Per aconseguir aquests objectius, s'han utilitzat una àmplia varietat de tecnologies i estratègies que han anat canviant amb el temps, reflectint el desenvolupament dels mètodes disponibles per a l'estudi de virus (Hood et al., 2021). Cada una d'aquestes estratègies, individualment, pot aportar informació útil, i sovint són necessàries combinacions d'aquestes per proporcionar avaluacions de risc més completes (Santarpia et al., 2023).

2.1. Mètodes de mostreig i processament de mostres ambientals

A l'ambient, les potencials fonts de virus humans i, per tant, d'infecció, inclouen l'aigua, les aigües residuals i els fangs de depuradora, els fems, l'aire, les superfícies i els fòmits, les fruites i verdures i alguns animals, com els mol·luscs bivalves (Rodríguez-Lázaro et al., 2012). Per tal de desenvolupar un pla de mostreig per estudiar la presència de virus en aquestes matrius ambientals cal definir l'objectiu, els perills potencials i les accions a seguir que permetin seleccionar les estratègies i les tecnologies més adequades i que garanteixin l'obtenció de les dades desitjades (Figura 5). Conèixer la transmissió del virus o dels virus que es volen estudiar aporta informació rellevant per al disseny del mostreig i la selecció del tipus de mostra més apropiada (cercle interior de la Figura 5).

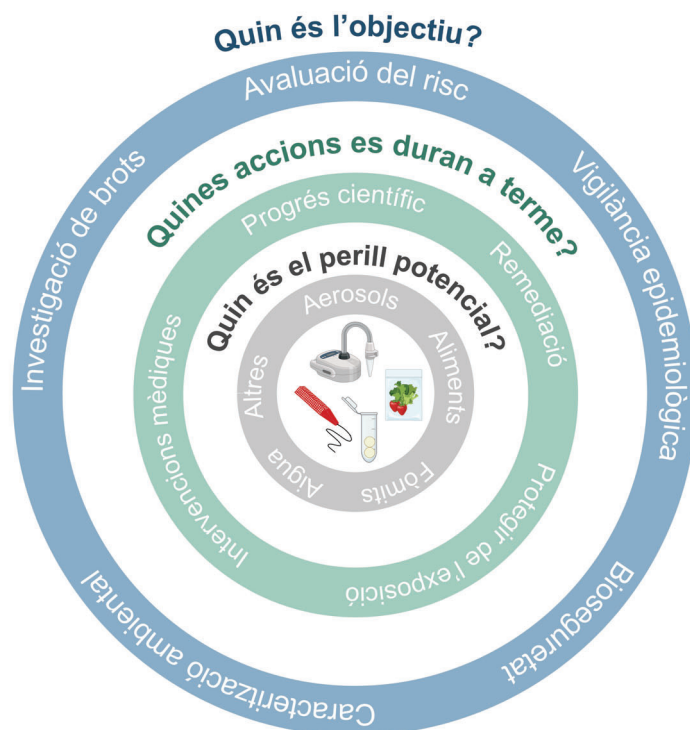


Figura 5. Marc per al desenvolupament d'un pla de mostreig. Les tres preguntes clau i les seves respostes generals es situen a l'interior dels cercles, mentre que les icones fan referència a algunes de les tecnologies emprades per al mostreig d'aerosols, aliments, fòmits i aigua. Adaptada de Santarpia et al., 2023

Els mètodes de mostreig i de processament de mostres ambientals difereixen en funció de les característiques de la matriu i dels agents vírics a estudiar, així com de les posteriors anàlisis que es vulguin realitzar (Figura 6). De forma general, després del mostreig específic per cada tipus de mostra ambiental, les partícules víriques han de ser separades de la matriu i, si és necessari, concentrades a un volum adequat per poder ser detectades posteriorment. Si la detecció té com a diana els àcids nucleics, caldrà realitzar un pas d'extracció d'aquests previ a la seva detecció.

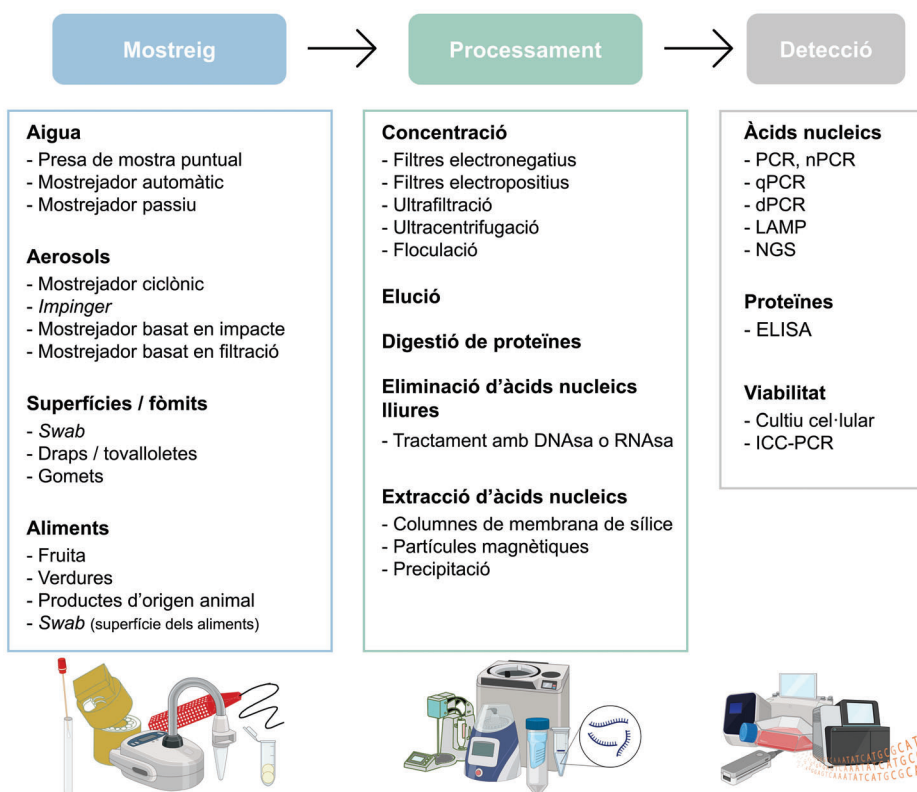


Figura 6. Flux general per a la detecció de virus en mostres ambientals, amb els principals mètodes de mostreig en funció de la matriu analitzada i passos a tenir en compte durant el seu processament per tal de poder dur a terme la detecció posterior.

Mostreig

Els mostreigs d'aigües normalment es basen en l'obtenció de mostres puntuals o bé mostres compostes mitjançant l'ús de mostrejadors automàtics. Aquests últims sobretot s'utilitzen en estudis d'epidemiologia basada en aigües residuals, ja que recullen mostres compostes proporcionals al flux diari i/o a un interval de temps determinat, obtenint així mostres més representatives (W. Ahmed et al., 2022). L'ús de mostrejadors passius s'ha emprat com una alternativa econòmica i pràctica, basant-se en la instal·lació d'un dispositiu, amb una membrana o material que permeti la retenció de virus, durant un període de temps

determinat. Un exemple és la unitat de mostreig amb aparença de torpede i impresa en 3D, dissenyada per Schang i col·laboradors (Schang et al., 2021). Aquesta unitat s'ha utilitzat amb diferents tipus de materials per retenir virus presents en l'aigua residual, essent les membranes electronegatives les que s'han reportat com el material més adequat per a ser col·locat dins dels torpedes (Mejías-Molina et al., 2023).

El mostreig de bioaerosols és una àrea d'investigació activa, ja que aquests suposen un repte per la seva gran diversitat en termes de mida, propietats biològiques i requisits per a la seva detecció i quantificació. Per al mostreig de virus en aire, existeixen diferents tipus de mostrejadors de bioaerosols disponibles comercialment, entre els quals es troben els ciclònics, els *impingers* i els basats en impactació o filtració (Mainelis, 2020). L'elecció del mostrejador més adequat es basa en molts factors, incloent-hi la disponibilitat, el cost, el volum d'aire que es vulgui mostrejar, la mobilitat, l'eficiència i les condicions ambientals en les quals es durà a terme l'obtenció de mostres (Pepper & Gerba, 2015).

L'eina més utilitzada per al mostreig de virus en superfícies és l'hisop (*swab*), que pot ser de diferents materials. No obstant això, també s'han utilitzat altres tècniques, com ara draps i gases i, més recentment, s'ha descrit l'ús d'adhesius de paper, coneguts popularment com a "gomets", i que suposa una nova estratègia per mostrejar virus en superfícies a llarg termini (Bobal et al., 2019; Sommer et al., 2021).

Pel que fa als aliments, existeixen metodologies estandarditzades disponibles associades a la detecció i quantificació de HAV i NoV, per a un nombre limitat de matrius, incloent-hi fruites petites sense llavors (maduixes, gerds o groselles), verdures de fulla, tija i bulb, mol·luscs bivalves i superfícies d'aliments (ISO, 2017). En aquestes, s'especifica que les fruites i verdures a analitzar han de ser fresques o congelades, els mol·luscs bivalves han d'estar vius o, si estan congelats, intactes i sense

haver passat cap tractament tèrmic. Si el mostreig es fa de la superfície dels aliments, s'indica que s'ha d'utilitzar un hisop de cotó estèril prèviament humitejat amb tampó fosfat salí (PBS).

Processament

Un cop obtingudes les mostres, el processament d'aquestes pot ser molt variat en funció del tipus de matriu i l'objectiu final de l'estudi. Un pas que sovint és necessari en mostres ambientals és la concentració de partícules víriques, ja que els virus habitualment es troben en baixa quantitat i, sense aquest pas, no es podrien detectar (Stoufer et al., 2024). En el cas de la concentració de virus en aigües, encara que s'utilitzen alguns mètodes de concentració d'un sol pas, la majoria dels mètodes consisteixen en dos passos, permetent així la concentració de partícules víriques presents en el volum de mostra inicial, que pot ser de pocs fins a centenars de litres, a uns pocs mil·lilitres. D'entre tots els mètodes de concentració vírica disponibles, el tipus i el volum d'aigua a processar són factors essencials en la selecció del mètode més adequat, ja que no hi ha un mètode únic per a tots els tipus de mostres (Bofill-Mas & Rusiñol, 2020). En aigües residuals, els virus es troben en concentracions més elevades i, per tant, són detectables fàcilment amb volums de mostreig relativament petits (<100 ml). En canvi, la detecció de virus en aigües superficials o subterrànies requereix volums més grans de mostreig (10 – 100 L), ja que es troben en menor concentració, excepte si hi ha hagut episodis puntuals de contaminació com ara fuites de clavegueram o situacions de pluja torrencial (Rusiñol et al., 2015). Actualment, una àmplia varietat de mètodes es basa en diversos processos fisicoquímics, que impliquen l'adsorció de partícules víriques en diferents matrius o la seva retenció segons la seva mida (Bofill-Mas & Rusiñol, 2020).

Altres possibles passos a realitzar durant el processament de mostres ambientals poden ser l'elució, que consisteix en el rentat amb *buffers* apropiats que permeten alliberar les partícules víriques de la matriu, així com els tractaments amb DNAsa o RNAsa o l'extracció d'àcids nucleics, si posteriorment es vol realitzar un mètode de detecció que tingui aquesta diana.

Tal i com hem vist, les metodologies estandarditzades disponibles són per a un nombre limitat de tipus de matrius (superfícies d'aliments, tipus concrets de fruita, verdures, mol·luscs bivalves i aigua embotellada) associades a la detecció de HAV i NoV (ISO, 2017). Tot i que són recursos útils, no s'apliquen a tota la diversitat de matrius ambientals existent ni estan validats per a altres virus, i sovint cal fer servir el millor criteri i els recursos disponibles per determinar quin mètode és l'adequat per a l'anàlisi dels virus en les mostres que es volen estudiar (Stoufer et al., 2024). Per aquest motiu, l'establiment de mètodes estàndard que informin sobre les taxes de recuperació víriques esperades per cada tipus de matriu ambiental encara segueix sent una necessitat important (Wang et al., 2023).

2.2. Eines tradicionals de detecció de virus en mostres ambientals

Històricament, els mètodes de cultiu cel·lular eren els més utilitzats per a la detecció de virus infecciosos (Haramoto et al., 2018). Aquests es basen en l'observació de l'efecte citopàtic, és a dir, els canvis que genera la infecció vírica en la cèl·lula hoste. No obstant, els principals desavantatges del cultiu cel·lular són l'existència de virus que no produeixen aquest efecte de manera evident o el fet que no es disposin de línies cel·lulars hoste per a molts virus d'interès (Wang et al., 2023). Des de mitjans de la dècada dels anys 80, el desenvolupament de les

tècniques basades en biologia molecular va suposar un gran avanç de la virologia ambiental gràcies al fet que la reacció en cadena de la polimerasa (PCR) va esdevenir una tècnica de referència. Aquesta tècnica és més ràpida, sensible, específica i econòmica que els mètodes de detecció mitjançant cultiu cel·lular i, a més, permet la detecció de virus que no són cultivables. La PCR, juntament amb altres estratègies també basades en l'amplificació d'àcids nucleics, com l'amplificació isotèrmica mediada per bucle (LAMP), han millorat considerablement la sensibilitat de detecció i han permès la detecció en matrius ambientals amb baixa concentració de virus (Yao et al., 2021). Aquestes tècniques permeten la detecció de virus mitjançant l'ús d'encebadors específics que amplifiquen regions d'interès del genoma víric. En cas que es vulgui detectar un virus amb genoma ARN, cal fer un pas de retrotranscripció (RT) previ a la reacció d'amplificació. Degut a que la concentració de virus en mostres ambientals sol ser baixa, sovint es requereix d'una segona PCR amb encebadors més interns anomenada PCR niuada (nPCR), augmentant així la sensibilitat del mètode.

La PCR a temps real, també coneguda com a PCR quantitativa (qPCR) va representar el progrés més significatiu en la detecció de virus en entorns aquàtics (Hamza & Bibby, 2019). L'ús d'una sonda marcada amb un fluoròfor permet monitoritzar el procés d'amplificació dels àcids nucleics a temps real en instruments especialitzats capaços de recollir les dades de fluorescència després de cada cicle de PCR, permetent així la quantificació de la càrrega vírica present en una mostra. A més, aquesta tècnica es pot dissenyar per amplificar múltiples dianes a la vegada, de manera que possibilita detectar i quantificar diferents virus de forma simultània en un mateix assaig (Yao et al., 2021).

La PCR digital (dPCR) ha guanyat l'atenció en els últims anys com a mètode per quantificar de manera precisa i absoluta, sense necessitat de dependre d'una corba estàndard com en els assaigs de qPCR. La dPCR

es basa en dividir aleatòriament la mostra en reaccions separades, com ara en milers de gotes, teòricament amb cap o una molècula d'àcid nucleic diana present en cada cambra de reacció. A continuació, cada cambra s'utilitza per realitzar una amplificació independent, i la quantificació absoluta s'estima amb estadístiques de Poisson. Aquest sistema permet que la dPCR proporcioni una quantificació absoluta altament sensible i precisa, reduint els falsos negatius i millorant la quantificació de virus presents en mostres ambientals complexes (Yao et al., 2021), ja que aquest mètode es veu menys afectat pels inhibidors de PCR. S'han desenvolupat diferents plataformes de dPCR, com per exemple la dPCR basada en gotes (ddPCR), que combina la tecnologia d'emulsió de gotes d'aigua-oli amb la microfluídica (Haramoto et al., 2018).

Existeixen recomanacions per reportar resultats de qPCR o dPCR, com ara les MIQE (Bustin et al., 2009), que són directrius sobre la informació mínima per a la publicació d'experiments qPCR, o les EMMI (Borchardt et al., 2021), que són directrius sobre la informació mínima per reportar dades de qPCR i dPCR en estudis de microbiologia ambiental. Aquestes recomanacions aporten informació detallada sobre els tipus de controls i la informació que hauria d'acompanyar aquestes dades per garantir que els estudis tinguin resultats comparables i reproduïbles.

Un desavantatge important dels mètodes moleculars és que no permeten diferenciar els virus infecciosos dels que no ho són. Per tal de superar aquest inconvenient, s'han desenvolupat estratègies com el cultiu cel·lular integrat a PCR (ICC-PCR), una combinació de cultiu cel·lular i molecular que incorpora els avantatges de tots dos mètodes per superar les mancances de cadascun. També s'han ideat altres estratègies independents de cultiu per distingir les partícules víriques potencialment infeccioses, com a pas previ a l'extracció d'àcids nucleics, com l'ús de nucleases que eliminen l'ADN o ARN lliure, el tractament enzimàtic de les

partícules víriques, l'ús d'intercalants i la captura immunomagnètica de partícules víriques (Haramoto et al., 2018; Rodríguez et al., 2009).

Altres mètodes tradicionals de detecció de virus són la identificació mitjançant microscòpia electrònica o també els immunoassaigs, els quals es basen en la interacció altament específica entre els anticossos i els seus antígens, com per exemple l'assaig d'immunoadsorció lligat a enzims (ELISA). No obstant, la seva sensibilitat analítica és baixa per a una detecció eficaç en mostres ambientals (Rodríguez-Lázaro et al., 2012).

La pròpia naturalesa dels mètodes tradicionals de detecció de virus implica que moltes vegades no es pugui reflectir la diversitat vírica real que hi ha en una mostra, ja que són mètodes específics o estan enfocats a un grup determinat de virus, requerint conèixer prèviament les propietats del virus d'estudi (Kumar et al., 2017). En canvi, els avanços amb les tecnologies de seqüenciació han implicat la transformació del camp del descobriment dels virus i l'anàlisi dels viromes.

2.3. Noves metodologies: seqüenciació massiva per a la detecció de virus en mostres ambientals

El desenvolupament de la metagenòmica amb les tècniques de seqüenciació de nova generació (NGS) ha canviat radicalment el món de la virologia, permetent la descripció de les comunitats víriques, el viroma, en diferents ecosistemes sense haver de tenir un coneixement preconcebut de la seva composició. A més, ha facilitat la detecció, identificació, i caracterització dels virus que constitueixen el viroma humà (Geng-Hao et al., 2022). També ha suposat afrontar nous reptes, com ara el fet d'obtenir una gran quantitat de seqüències víriques però sovint amb informació limitada sobre els seus hostes o altres propietats biològiques (Simmonds et al., 2017). De fet, molts estudis han utilitzat el terme

“matèria obscura” (*“dark matter”*) per referir-se a la gran quantitat de seqüències obtingudes mitjançant la seqüenciació massiva, la majoria de les quals probablement correspon a virus encara desconeguts.

El procediment habitual que es segueix per a estudiar el viroma complet en mostres ambientals mitjançant NGS es detalla a la Figura 7. La limitació que suposa fer aquest tipus d'estudis és que els genomes vírics en mostres ambientals solen trobar-se en menor proporció respecte d'altres genomes, com ara bacterians, motiu pel qual es requereix dur a terme una sèrie de procediments previs a la seqüenciació per tal de reduir els àcids nucleics no vírics, incloent-hi mètodes de concentració i d'extracció que ho afavoreixin, així com l'ús de nucleases per eliminar l'ADN/ARN lliure.

A més, sovint també és necessària una pre-amplificació dels àcids nucleics vírics presents en la mostra mitjançant una PCR amb encebadors aleatoris, com per exemple l'amplificació per desplaçament múltiple (MDA) o l'amplificació independent de seqüència basada en únic encebador (SISPA). L'estudi del viroma d'una mostra requereix d'aquests encebadors aleatoris perquè no existeix cap element genètic compartit per tots els virus, a diferència dels bacteris, la qual cosa dificulta l'estudi de la diversitat vírica. No obstant, sí que s'han descrit gens compartits entre els membres d'un mateix grup taxonòmic de virus que sí que permeten estudiar la diversitat dins d'aquest grup en concret (Breitbart & Rohwer, 2005). Aquest tipus de seqüenciació dirigida a un grup víric concret es detalla a l'apartat 2.3.2.

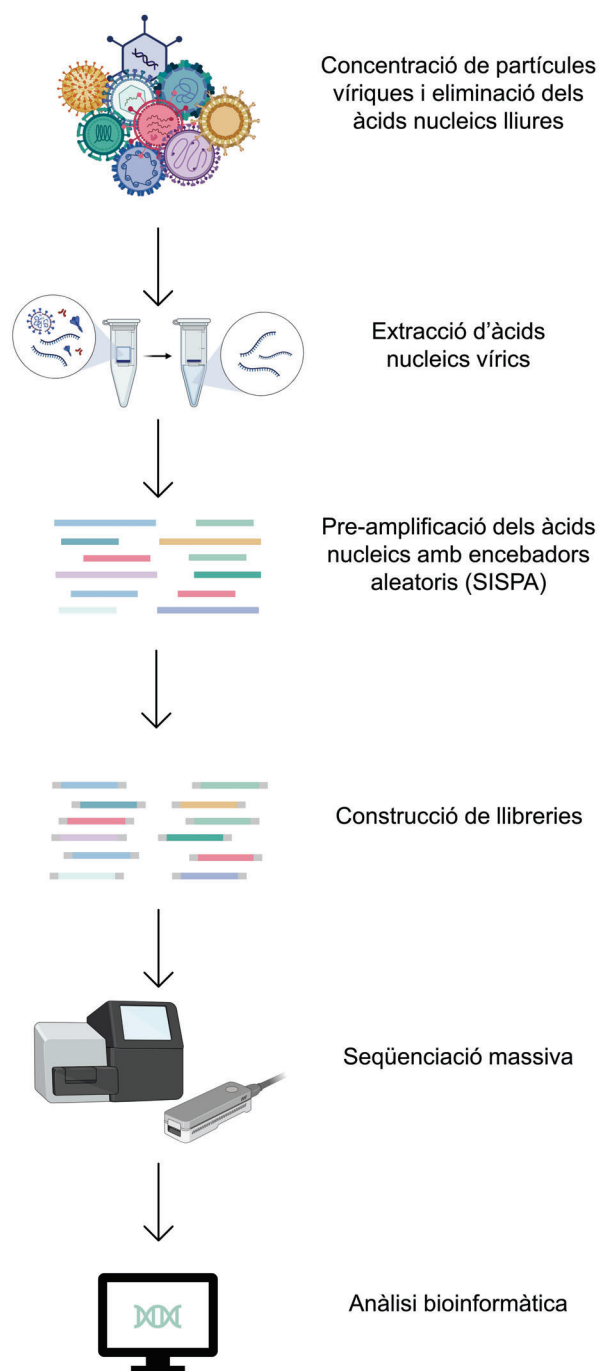


Figura 7. Flux de treball general per estudiar el viroma present en mostres ambientals mitjançant NGS.

Per tant, l'amplificació amb encebadors aleatoris permet l'estudi del viroma complet d'una mostra sense dirigir la seqüenciació a cap família o grup taxonòmic de virus concret. Aquest tipus de seqüenciació és el que es coneix com a seqüenciació no dirigida o d'alt rendiment (HTS), i té com a objectiu la seqüenciació de tots els genomes vírics presents en una mostra. Un cop feta aquesta pre-amplificació i haver incrementat la concentració d'ADN víric, es construeixen les llibreries de seqüenciació. Existeix una gran diversitat de protocols per a la preparació de llibreries de seqüenciació, que varien en funció de la plataforma de seqüenciació utilitzada o de la seva automatització (Head et al., 2014). Generalment, aquests protocols comencen amb una fragmentació de l'ADN, que pot ser física o enzimàtica. No obstant, aquest pas no és necessari en plataformes que permetin la seqüenciació de fragments llargs, com és el cas de les plataformes de seqüenciació desenvolupades per *Oxford Nanopore Technologies*. Posteriorment, es duu a terme una reparació dels extrems de les seqüències obtingudes i una lligació dels adaptadors de seqüenciació, que són necessaris per al reconeixement i la unió al suport de la plataforma de seqüenciació. A més, si es volen seqüenciar diverses mostres alhora, es pot efectuar també la lligació d'índexs, que són uns marcadors específics que permeten etiquetar les seqüències i, d'aquesta manera, poder identificar les mostres. Una vegada s'han lligat els índexs i els adaptadors, es pot fer un pas de selecció de fragments per mida i, finalment, segons el protocol utilitzat, les llibreries poden ser amplificades per PCR i purificades abans de ser carregades a la plataforma de seqüenciació (Martínez-Puchol, 2020b).

Finalment, les dades obtingudes amb la seqüenciació són processades i analitzades bioinformàticament. Breument, aquest processament consisteix en un filtrat per eliminar els *reads* de baixa qualitat, i els restants són assemblats per formar seqüències de major longitud, anomenades *contigs*. Després, es realitza l'assignació taxonòmica de les seqüències

processades mitjançant l'alineament amb una base de dades, com ara GenBank (NCBI, 2013) o UniProt (Bateman, 2019). Donada la complexitat d'aquest procediment d'anàlisi bioinformàtica, han sorgit eines específicament dissenyades per a l'anàlisi de seqüències víriques i que permeten un tractament de les dades de manera fàcil, ràpida i accessible per a usuaris que no siguin especialistes en bioinformàtica. Alguns exemples d'eines web que permeten el processament de les dades generades per les plataformes de NGS de forma gratuïta són *Genome Detective* (<https://www.genomedetective.com>) (Vilsker et al., 2019) i *CZ ID* (<https://czid.org/>) (Kalantar et al., 2020).

Considerant que les eines de NGS poden proporcionar informació fonamental sobre les seqüències d'àcids nucleics víriques, com per exemple l'origen, l'evolució i les mutacions, aquestes han demostrat ser una eina valuosa per a la detecció i el descobriment de virus, així com per als estudis de diversitat vírica (Yao et al., 2021). L'aplicació d'aquestes eines ha fet possible l'estudi de virus en un ampli ventall de matrius ambientals, com ara aigües residuals (Cantalupo et al., 2011; Fernandez-Cassi et al., 2018; Martínez-Puchol et al., 2020a), oceans (Brum et al., 2015; Hurwitz et al., 2013), llacs (Djikeng et al., 2009), aigües subterrànies (Kallies et al., 2019; Zaouri et al., 2020), aigües de reg (Fernandez-Cassi et al., 2017; Rusiñol et al., 2020), aliments (Aw et al., 2016; Fernandez-Cassi et al., 2017; Zhang et al., 2014), aerosols (Hall et al., 2013; Han et al., 2019), superfícies de transport públic (Hsu et al., 2016) i, fins i tot, telèfons mòbils, fòmits, de treballadors de l'àmbit sanitari (Boucherabine et al., 2022). En aquests estudis sobre els virus presents en mostres ambientals s'ha utilitzat majoritàriament la seqüenciació HTS, la qual permet l'estudi del viroma general sense dirigir la seqüenciació cap a famílies víriques o grups de virus concrets. No obstant, en mostres complexes on les seqüències víriques poden representar només un 1% de les dades generades, aquesta aproximació pot suposar la pèrdua de

seqüències víriques d'interès que es trobin en menor proporció (Martínez-Puchol, 2020b). És per això que s'han desenvolupat tècniques de seqüenciació massiva amb enriquiment de dianes o també mitjançant la generació d'amplicons, que permeten la detecció de famílies o grups vírics d'interès (Figura 8).

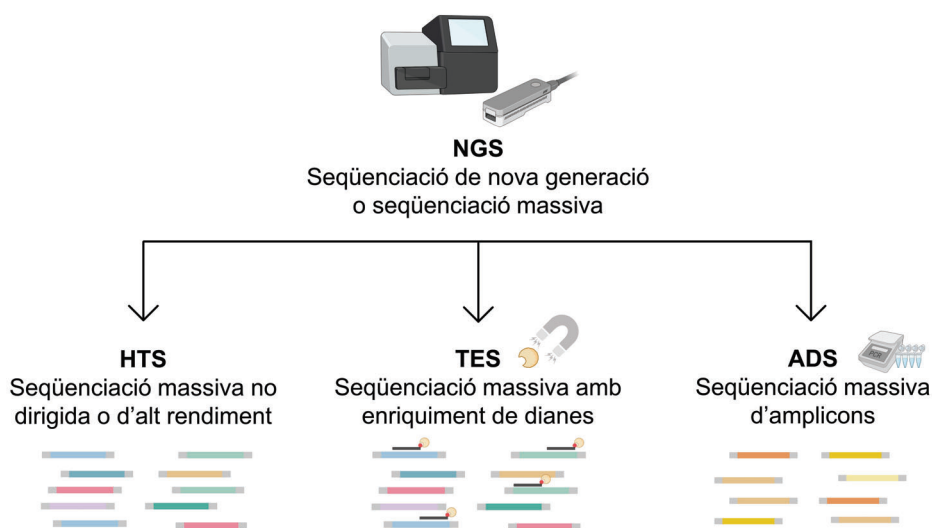


Figura 8. Diferents tècniques de seqüenciació massiva o NGS.

2.3.1. Seqüenciació massiva amb enriquiment de dianes

Una estratègia per tal de millorar la seqüenciació de determinades famílies o grups vírics és la seqüenciació massiva amb enriquiment de dianes (TES). L'objectiu d'aquest enriquiment és dirigir la seqüenciació cap als virus d'interès i no perdre rendiment amb seqüències que no estiguin dins de l'àmbit d'estudi, augmentant així la sensibilitat i eficiència de detecció.

El panell de captura VirCapSeq-VERT (Roche), el qual consisteix en aproximadament 2 milions de sondes que cobreixen els genomes de 207 virus coneguts que infecten vertebrats, ha demostrat ser útil per a la detecció de seqüències víriques en mostres complexes (Briese et al.,

2015). No obstant, és limitat el nombre d'estudis que l'han utilitzat en matrius ambientals, incloent-hi aigües residuals (Hjelmsø et al., 2019; Martínez-Puchol et al., 2020a; Mejías-Molina et al., 2023) i ostres (Strubbia et al., 2020).

La Figura 9 representa de forma gràfica la seqüenciació massiva amb enriquiment de dianes mitjançant la hibridació amb sondes. Durant la preparació de la llibreria de seqüenciació, s'introdueix el pas d'enriquiment en què es produeix la hibridació d'aquestes sondes, conjugades amb biotina, amb les seqüències víriques d'interès. Posteriorment a aquesta hibridació, s'utilitzen unes boles magnètiques recobertes amb estreptavidina que permeten la captura dels híbrids i l'eliminació de les seqüències que no siguin diana. Finalment, la llibreria capturada és amplificada, obtenint una major profunditat de seqüenciació i proporcionant una comprensió més completa de la diversitat vírica que es vol estudiar, a més de permetre el descobriment de potencials virus emergents (Pogka et al., 2022).

Existeixen altres panells de captura de diferents cases comercials, com per exemple *Illumina* o *Twist Bioscience*, dirigits a l'enriquiment de seqüències de grups de virus concrets, com ara panells de captura de virus respiratoris, de diferents gèneres de coronavirus o dedicats exclusivament a SARS-CoV-2 (Ceballos-Garzon et al., 2024). Actualment, inclús s'ofereix la possibilitat de dissenyar aquests panells de forma personalitzada, de manera que l'usuari pot decidir quines sondes incloure al panell i dirigir així la seqüenciació cap a aquells virus concrets que es vulguin estudiar.

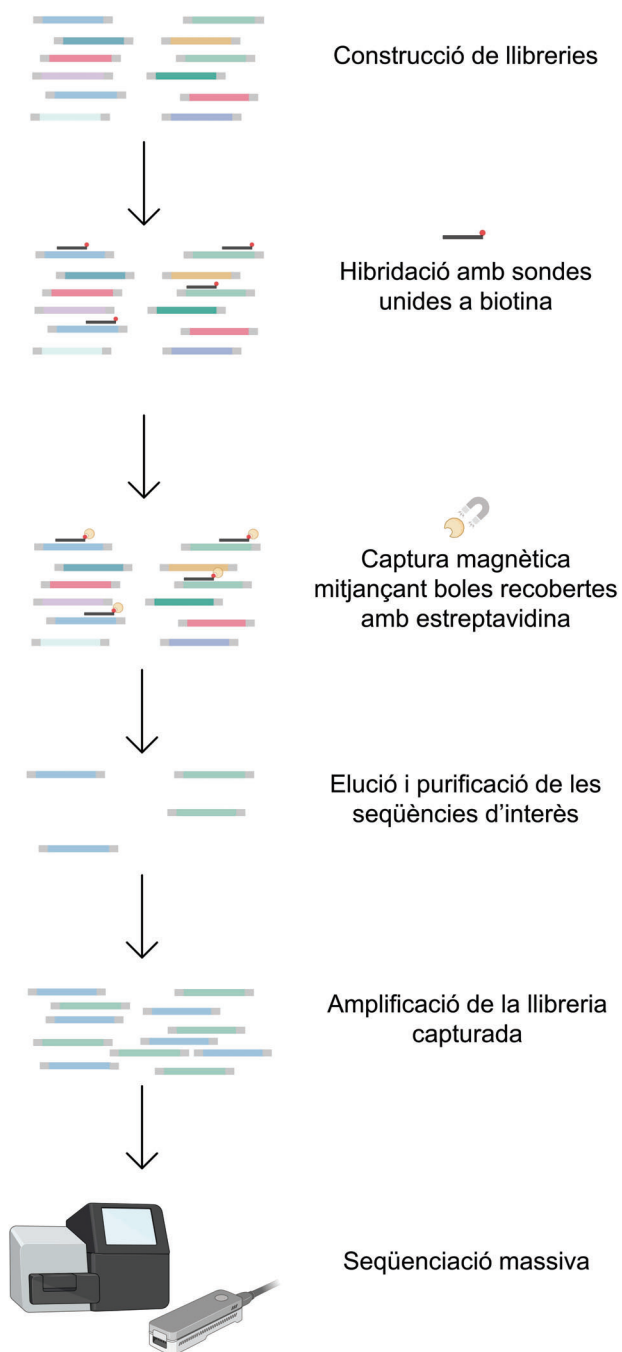


Figura 9. Seqüenciació massiva amb enriquiment de dianes mitjançant la hibridació amb sondes. Adaptada de Martínez-Puchol, 2020b.

La seqüenciació massiva amb enriquiment de dianes està en constant desenvolupament. La principal limitació que pot suposar l'enriquiment de seqüències amb aquests panells és que es produeixi un biaix cap a seqüències de virus coneguts. No obstant, fer menys restrictiva la hibridació pot possibilitar el descobriment de virus nous. L'altra limitació que pot suposar la hibridació de sondes és la falta de cobertura del genoma víric, la qual cosa podria dificultar l'estudi de variants si les regions clau per tipificar no hibriden amb les sondes. En aquests casos, la seqüenciació massiva d'amplicons pot ser una bona alternativa com a metodologia per a la caracterització de virus dins d'un mateix grup o família vírica (Martínez-Puchol, 2020b).

2.3.2. Seqüenciació massiva d'amplicons

La seqüenciació d'amplicons obtinguts per PCR s'ha realitzat tradicionalment mitjançant el mètode Sanger (Sanger et al., 1977), però amb aquest mètode només s'obté la seqüència predominant i no mostra la diversitat de seqüències present en una mostra. En canvi, amb la seqüenciació massiva dels amplicons (ADS) sí que s'obté el perfil de la diversitat de seqüències amplificades amb la PCR (Figura 10), permetent així l'estudi detallat de famílies específiques.

És una aproximació relativament senzilla i sensible, però existeixen certes limitacions com ara la longitud marcada per la reacció de PCR, així com la localització de regions on puguin unir-se els encebadors alhora que la regió amplificada sigui prou específica per poder diferenciar membres dins d'un mateix grup víric (Kuchinski et al., 2022).

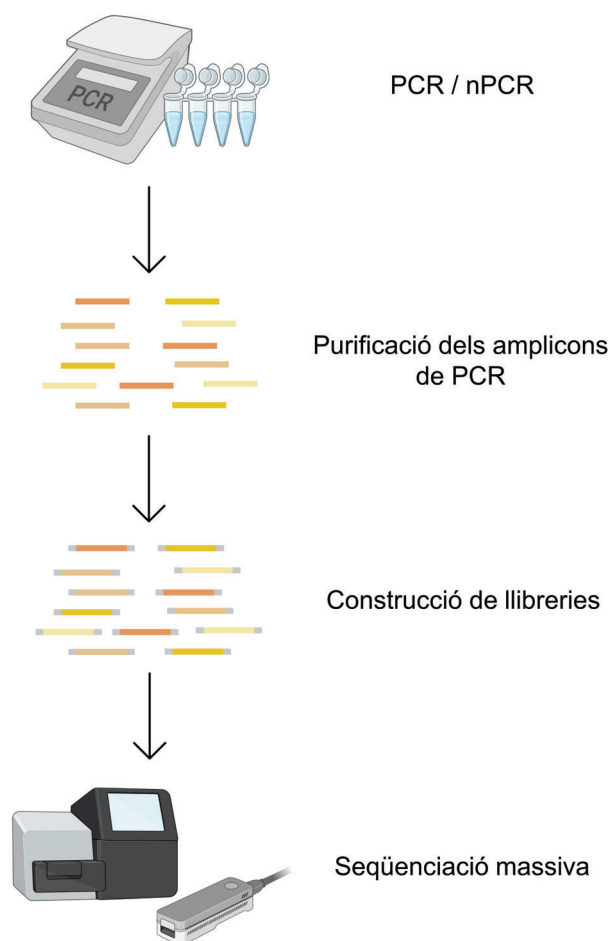


Figura 10. Seqüenciació massiva d'amplicons. Adaptada de Martínez-Puchol, 2020b.

No obstant, aquesta aproximació ha demostrat ser una eina útil per als estudis de diversitat vírica dins d'una mateixa família i permet detectar variants minoritàries (Martínez-Puchol et al., 2020b; Radford et al., 2012). Aquesta estratègia s'ha utilitzat per a l'estudi de determinats grups de virus en mostres ambientals, com ara HAdV en aigües residuals i rius (Fernandez-Cassi et al., 2018; Martínez-Puchol et al., 2020a; Ogorzaly et al., 2015), papil·lomavirus humans (HPV) en aigües residuals (Hamza & Hamza, 2018; Martínez-Puchol et al., 2020a), rius (Iaconelli et al., 2015) i aigües de piscina (La Rosa et al., 2015), i NoV en aigües residuals

(Martínez-Puchol et al., 2020a; Prevost et al., 2015) i mol·luscs bivalves (Desdouits et al., 2020; Imamura et al., 2017; Oshiki et al., 2018). A més, s'ha demostrat el seu potencial com a estratègia per a la vigilància de virus patògens específics, incloent-hi la identificació i caracterització de les variants de SARS-CoV-2 en aigües residuals (Bar-Or et al., 2021; Barbé et al., 2022; Segelhurst et al., 2023). El protocol ARTIC (<https://artic.network>), d'accés obert, ha estat àmpliament utilitzat amb aquesta finalitat, i s'ha dissenyat per amplificar el genoma complet del virus ja que consisteix amb una PCR múltiplex de 98 parells d'encebadors (ARTIC Network, 2020).

3. Avaluació del risc microbiològic associat a l'exposició a virus

El ràpid desenvolupament de mètodes de detecció de microorganismes patògens en el nostre entorn en les últimes dècades ha generat un augment en l'interès per desenvolupar eines d'avaluació dels riscos per a la salut associats a aquests (Nwachuku & Gerba, 2004). Tot i que les malalties infeccioses han existit des dels inicis de la humanitat, no va ser fins l'arribada dels mètodes de detecció que els patògens van ser identificats com a perill i que l'ambient va ser reconegut com a via d'exposició (Haas et al., 2014).

Concretament, l'avaluació del risc consisteix en la caracterització qualitativa o quantitativa i l'estimació dels possibles efectes adversos per a la salut associats amb l'exposició d'individus o poblacions a un determinat perill (Haas et al., 2014), en aquest cas microbiològic. És una avaluació precisa per estimar els riscos, ja que organitza i analitza sistemàticament tota la informació científica i tècnica disponible, alhora que considera les variabilitats i les incerteses (Van der Fels-Klerx et al., 2018). No obstant, l'avaluació del risc no s'utilitza de manera aïllada, sinó que forma part del que es coneix, en un context més ampli, com a anàlisi del risc.

L'anàlisi del risc és un procés utilitzat en diversos camps, incloent-hi negocis, assegurances i desastres naturals, però també en l'àmbit de la contaminació ambiental i la salut. De forma general, l'anàlisi de risc consta de tres components interconnectats: l'avaluació, la gestió i la comunicació del risc (Figura 11). En primer lloc, és important identificar i caracteritzar el perill potencial per a un individu o una població, així com avaluar-ne l'exposició i els seus efectes sobre la salut. L'objectiu principal és la seguretat i, per tant, la gestió i la comunicació del risc són part integral d'aquest procés d'anàlisi.

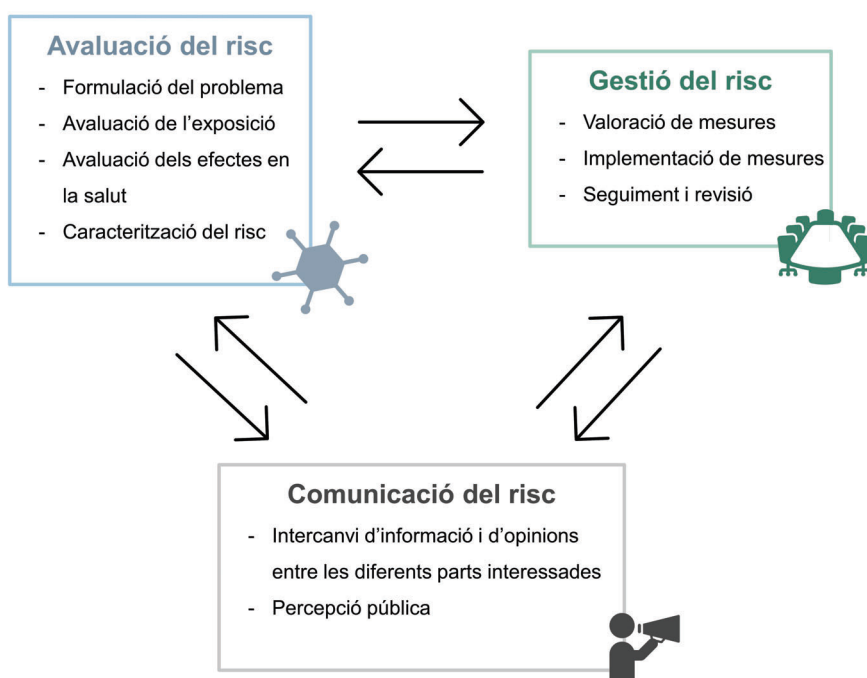


Figura 11. Estructura i components de l'anàlisi de risc. Adaptada de Fazil, 2005.

L'avaluació qualitativa del risc permet determinar si un perill concret necessita una anàlisi addicional i, a diferència de la quantitativa, permet categoritzar els riscos tot i tenir grans buits de dades. La probabilitat del resultat o la magnitud de les conseqüències s'expressa en termes com "alt", "mitjà" o "baix" (Macdiarmid & Pharo, 2003). En canvi, amb l'avaluació quantitativa es pretén determinar numèricament les probabilitats i, per tant, requereix una major exigència de dades i una comprensió més exhaustiva del risc.

3.1. Avaluació quantitativa del risc microbiològic

L'avaluació quantitativa del risc microbiològic (QMRA) és una eina teòrica que permet estimar la probabilitat d'una infecció o d'una malaltia a causa de l'exposició d'un individu a un agent microbiològic infecciós. Aquesta eina s'ha utilitzat durant més de 35 anys (Haas, 1983a, 1983b) i ha resultat ser útil en diversos estudis sobre exposició a virus patògens en diferents matrius ambientals, com ara bioaerosols (Brooks et al., 2005; Brooks et al., 2012; Carducci et al., 2018; Dada & Gyawali, 2021; Medema et al., 2004; Paccha et al., 2016; Tanner et al., 2008), superfícies o fòmits (Brooks et al., 2012; Lanzarini et al., 2022; Paccha et al., 2016), aigua de beguda i aliments (Brooks et al., 2012; Deere & Ryan, 2022; Gonzales-Gustavson et al., 2019; Petterson, 2016; Sales-Ortells et al., 2015b; Schijven et al., 2019), aigües regenerades i de reg (Antwi-Agyei et al., 2016; Schoen et al., 2018), aigües recreatives (Girardi et al., 2019; Gularte et al., 2019; Sales-Ortells et al., 2015a; Sales-Ortells & Medema, 2014) i aigües residuals (Zaneti et al., 2021).

Per la seva pròpia naturalesa, la QMRA és interdisciplinària i, per tant, requereix la participació de diverses disciplines establertes, incloent-hi l'economia, l'enginyeria, les matemàtiques, les ciències polítiques i socials, així com les ciències naturals (Haas et al., 2014). Aquesta eina ha permès establir criteris per al tractament de l'aigua en diferents contextos (producció d'aigua potable i reutilització d'aigües), desinfecció de superfícies, processament d'aliments i purificació de l'aire (Blatchley & Cui, 2023; Dada & Gyawali, 2021; Haas, 2020).

L'aplicació de QMRA per a l'exposició ambiental a un virus patògen consisteix en quatre passos clau (Figura 12):



Figura 12. Marc de treball de la QMRA. Adaptada de WHO, 2016.

- 1. Formulació del problema.** Consisteix en identificar l'agent víric que pot estar present en una matriu ambiental concreta i que potencialment pot causar efectes adversos per a la salut. En aquest pas, també s'identifiquen les vies d'exposició i s'estableix la mètrica per expressar l'impacte que el virus té en la salut, com ara la probabilitat d'infecció i la probabilitat de malaltia, o bé un indicador que representi la càrrega de la malaltia agrupant tots els impactes.

- 2. Avaluació de l'exposició.** L'objectiu és estimar la magnitud i la freqüència de l'exposició al virus patògen a través de les vies d'exposició identificades durant la formulació del problema.
- 3. Avaluació dels efectes en la salut.** Consisteix en recopilar les dades sobre l'impacte del virus en la salut i la població d'estudi específica. L'avaluació és possible amb les relacions dosi-resposta identificades per cada virus patògen, que relacionen la dosi d'exposició amb la probabilitat d'infecció o malaltia, i la probabilitat de morbiditat i mortalitat. Altres factors rellevants per a aquesta avaluació són la transmissió secundària i la immunitat.
- 4. Caracterització del risc.** Es tracta de la integració del coneixement sobre l'agent víric i la seva exposició, amb combinació de l'avaluació dels efectes en la salut, per generar una mesura quantitativa del risc. D'aquesta manera, es pot predir la probabilitat d'un resultat advers en un individu o una població determinada, juntament amb les incerteses associades.

Aquestes fases permeten recopilar el coneixement científic sobre el patògen, el seu destí i el seu trànsit en l'ambient, així com les seves vies d'exposició i les malalties que causen com a conseqüència (Owens et al., 2020; WHO, 2016).

L'escala de temps en què s'expressa el risc pot variar, des d'esdeveniments d'exposició únics fins a totes les exposicions que es produeixen en un període determinat, com ara un any. A més, l'estimació del risc pot quantificar-se utilitzant diferents mètriques, incloent-hi la probabilitat d'infecció, la probabilitat de malaltia, el nombre esperat de casos de malaltia i mesures que representen la càrrega de la malaltia, com per exemple els anys de vida ajustats per discapacitat (DALYs). Aquesta última es calcula com la suma dels anys de vida perduts a causa

de la mort prematura en la població i els anys perduts per discapacitat per als casos incidents de la condició de salut (Deere & Ryan, 2022). Els DALYs van ser proposats per l'OMS com a mètrica per avaluar la càrrega de malaltia associada a exposicions ambientals, especialment a microorganismes, i s'ha utilitzat àmpliament per valorar les prioritats de salut pública (WHO, 2011). A més, en les directrius per a la qualitat de l'aigua de beguda, l'OMS va definir un límit de 10^{-6} DALY per persona i per any de càrrega de malaltia tolerable. Un altre límit de referència àmpliament utilitzat és el nivell acceptable de risc d'infecció anual proposat per l'Agència de Protecció Ambiental (EPA) dels Estats Units, el qual s'estableix a 10^{-4} per persona i per any (U.S. EPA, 2005). Ambdós límits de referència es van establir al voltant del objectiu de salut, fonamentats amb mètriques ben definides, i dels nivells de càrrega de malaltia tolerables (Fuhriemann et al., 2016). Determinar aquests límits per a un nivell de risc acceptable o tolerable és rellevant, ja que el risc zero és generalment reconegut com un objectiu inassolible.

L'aplicació de QMRA per a l'estimació de riscos vírics s'ha trobat amb diverses complicacions, com ara la baixa concentració de virus en l'ambient que dificulta la seva quantificació i tipificació. A més, no existeixen mètodes estandarditzats per a la quantificació de virus en diferents matrius ambientals i, en molts casos, no es disposen de models de dosi-resposta per als principals virus que es troben a l'ambient. Per obtenir quantificacions fiables, cal considerar diversos factors, com per exemple l'eficiència de detecció de l'assaig utilitzat, i s'han d'utilitzar paràmetres adequats per determinar amb precisió la concentració del virus a l'ambient (Rodríguez-Lázaro et al., 2012). Finalment, malgrat que s'aspira que l'anàlisi sigui la més objectiva possible, sovint manquen dades essencials i s'ha de recórrer a determinades assumpcions (Macdiarmid & Pharo, 2003). En aquest punt, és important que aquestes

assumpcions es declarin explícitament i es justifiquin adequadament per tal de preservar el rigor científic i transparència.

L'estudi dels virus presents en l'ambient i l'obtenció de dades adequades, són essencials, doncs, per a l'estimació dels riscos potencials que aquests poden representar per a la salut pública, i així poder implementar mesures de control adequades en funció de l'escenari de risc avaluat.

3.2. Risc ocupacional associat a l'exposició vírica

L'entorn laboral pot ser una font de disseminació vírica i, per tant, els treballadors poden estar exposats a determinats virus patògens que poden suposar un risc per a la seva salut. Les instal·lacions sanitàries són els entorns ocupacionals més estudiats, però en els últims anys també s'ha prestat atenció als treballadors relacionats amb el tractament d'aigües residuals, els quals poden estar exposats a virus de transmissió fecal-oral, així com personal veterinari que estigui en contacte amb animals i, per tant, pugui estar exposat a virus zoonòtics per manipulació de fems o per inhalació dels aerosols generats (Rodríguez-Lázaro et al., 2012).

L'exposició ocupacional als aerosols, en concret als bioaerosols, s'ha associat amb una àmplia diversitat d'efectes en la salut, incloent-hi malalties infeccioses, efectes tòxics aguts, al·lèrgies i càncer (Douwes et al., 2003). Els bioaerosols estan formats per petites partícules en suspensió que contenen microorganismes com ara virus, bacteris i fongs, i també compostos orgànics derivats d'aquests microorganismes (Mandal & Brandl, 2011). Els bioaerosols es poden generar a partir d'un gran nombre de processos i activitats dins l'entorn laboral, com són els tractaments d'aigües residuals, l'agricultura i la ramaderia (Mirskaya & Agranovski, 2018). Considerant que les aigües residuals contenen una gran diversitat de microorganismes patògens, entre els quals s'inclouen

virus com HAdV, RoV, HAV, NoV i Ast, la generació de bioaerosols durant el seu tractament pot afavorir l'alliberament d'aquests virus cap a l'atmosfera (Corpuz et al., 2020; Kataki et al., 2022). La gran majoria de les condicions de salut adverses reportades per treballadors d'estacions depuradores d'aigües residuals (EDARs) s'han atribuït als bioaerosols (Amoah et al., 2022; Carducci et al., 2016; Corrao et al., 2012).

Els bioaerosols generats i, per tant, els micoorganismes que els formen, poden acabar dipositant-se sobre les superfícies (Han et al., 2013). Per tant, els treballadors d'una EDAR no només poden estar exposats a virus mitjançant la inhalació de bioaerosols, sinó també a través de la ingestió oral després del contacte directe amb superfícies o fòmits contaminats, com ara les eines de treball. De fet, s'ha constatat que aquests treballadors tenen un major risc de desenvolupar símptomes en comparació amb la població general, incloent-hi símptomes respiratoris o gastrointestinals, relacionats amb l'entorn de treball (Masclaux et al., 2014).

No obstant, l'avaluació del risc ocupacional per a un treballador d'una EDAR és particularment difícil a causa de la gran diversitat de patògens presents en les aigües residuals, els diferents escenaris d'exposició, la naturalesa estocàstica de la infecció i els diferents nivells de susceptibilitat que poden tenir els treballadors, juntament amb la manca de dades d'estudis epidemiològics (Amoah et al., 2022). A més, la probabilitat d'infecció en aquests entorns depèn principalment de la capacitat d'infecció del patogen, la seva concentració i el temps d'exposició (Carducci et al., 2016).

L'emissió d'aerosols també s'ha convertit en una preocupació creixent en el sector de la producció porcina, especialment amb l'expansió ràpida de la seva activitat intensiva i sota la visió *One Health*. Els aerosols produïts a les granges de porcs principalment provenen de fonts com els fems, el

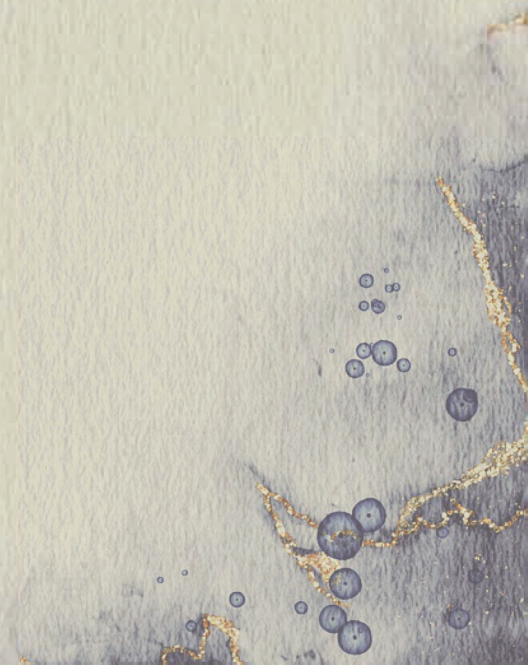
pinso, els pèls i la pell dels porcs, així com de la neteja i el tractament de residus (Liu et al., 2023). S'han detectat virus patògens a l'aire de granges de porc, com ara el virus de la grip A (IAV), el virus de la síndrome reproductiva i respiratòria porcina (PRRSV), el virus de la diarrea epidèmica porcina (PEDV), el virus de la pesta porcina clàssica (CSFV) i el virus de la pesta porcina africana (ASFV) (Alonso et al., 2015). A més, també s'ha demostrat el transport aeri de llarga distància de patògens com PRRSV i IAV (Corzo et al., 2013; Dee et al., 2009). Aquest últim, s'ha detectat també en superfícies d'instal·lacions d'explotació porcina (Neira et al., 2016). Tota aquesta informació suggereix certs riscos potencials per als treballadors d'aquestes instal·lacions, la qual cosa coincideix amb dades que indiquen que els treballadors del sector porcí tenen un major risc d'infectar-se amb un virus de la grip zoonòtic (Gray et al., 2007). És important, doncs, tenir en compte aquests treballadors per a la vigilància de pandèmies i per a les estratègies d'immunització (Myers et al., 2006).

Diversos estudis han aplicat QMRA per avaluar el risc d'infeccions víriques en diferents entorns laborals, concretament en EDARs (Carducci et al., 2016, 2018; Dada & Gyawali, 2021; Gholipour et al., 2021; Medema et al., 2004; Zaneti et al., 2021), zones agrícoles que utilitzen aigües residuals regenerades per a la irrigació (Antwi-Agyei et al., 2016) o aplicacions de fertilitzants (Brooks et al., 2005; Brooks et al., 2012; Tanner et al., 2008), instal·lacions d'explotació porcina (Paccha et al., 2016) i escenaris en què es manipulen lixiviats de residus sòlids municipals (Lanzarini et al., 2022).

L'aplicació de QMRA en estudis de risc ocupacional proporciona una base quantitativa per al procés de presa de decisions en la gestió d'aquest risc. Els models desenvolupats podrien ser útils per identificar les variables que es poden controlar des d'una perspectiva de gestió del risc, com ara el temps d'exposició i la concentració d'agents vírics, i establir intervencions i mesures de protecció, com la ventilació, tractaments de desinfecció o l'ús

d'equips de protecció individual (EPI) adequats, amb l'objectiu de reduir els riscos per sota dels límits establerts. No obstant això, ara per ara no s'han definit límits d'exposició ocupacional per a agents biològics. Per aquest motiu, l'ús de models de QMRA pot ser útil per a determinar aquests límits basant-se en un nivell acceptable de risc, permetent així l'elaboració de directrius per a la gestió del risc i conduint a una millor comprensió dels riscos per a la salut derivats de l'exposició a virus en entorns laborals (Carducci et al., 2016, 2018).

Objectius



L'objectiu global d'aquesta tesi és identificar els principals virus presents en l'ambient als quals els humans estem exposats i que constitueixen l'exposoma víric, així com avaluar el risc associat a l'exposició a virus en determinats entorns ocupacionals.

Els objectius específics han estat:

1. Optimitzar tècniques de NGS per a la detecció i caracterització de coronavirus en aigües residuals en un context de pandèmia.
2. Avaluar diferents mètodes per a la detecció i caracterització de variants de SARS-CoV-2 en aigües residuals.
3. Avaluar l'ús de la seqüenciació massiva amb enriquiment de dianes i la seqüenciació massiva d'amplicons per a la detecció i caracterització de patògens vírics en aliments de producció orgànica i en les seves fonts d'aigua de reg.
4. Identificar el viroma de les aigües subterrànies urbanes i de les seves potencials fonts de contaminació mitjançant la seqüenciació massiva amb enriquiment de dianes.
5. Avaluar l'ús d'adhesius de paper, "gomets", per a l'estudi de virus en superfícies per a l'obtenció de dades quantitatives sobre concentracions víriques necessàries per a anàlisis de QMRA.
6. Aplicar la seqüenciació massiva amb enriquiment de dianes per a l'estudi de virus presents en aire i superfícies.
7. Identificar virus humans i animals en entorns ocupacionals de risc i avaluar el risc microbiològic associat a l'exposició a aquests.

Informes dels articles



1. Llista d'articles inclosos a la tesi

Aquesta tesi està basada en les següents publicacions, enumerades segons l'ordre de presentació:

- **Itarte, M.**, Bofill-Mas, S., Martínez-Puchol, S., Torrell, H., Ceretó, A., Carrasco, M., Forés, E., Canela, N., Girones, R. & Rusiñol, M. (2021). Looking for a needle in a haystack. SARS-CoV-2 variant characterization in sewage. *Current Opinion in Environmental Science & Health*, 24, 100308.

<https://doi.org/10.1016/j.coesh.2021.100308>

- Martínez-Puchol, S., **Itarte, M.**, Rusiñol, M., Forés, E., Mejías-Molina, C., Andrés, C., Antón, A., Quer, J., Abril, J. F., Girones, R. & Bofill-Mas, S. (2021). Exploring the diversity of coronavirus in sewage during COVID-19 pandemic: Don't miss the forest for the trees. *Science of The Total Environment*. 8, 149562.

<https://doi.org/10.1016/j.scitotenv.2021.149562>

- Martínez-Puchol, S., Tarradas-Alemany, M., Mejías-Molina, C., **Itarte, M.**, Rusiñol, M., Baliellas, J., Abasolo, N., Canela, N., Monastiri, A., López-Roig, Serra-Cobo, J., Abril, J. F. & Bofill-Mas, S. (2024) Target enrichment metaviromics enables comprehensive surveillance of coronaviruses in environmental and animal samples. Article en procés de revisió a la revista *Heliyon*.

- **Itarte, M.**, Martínez-Puchol, S., Forés, E., Hundesa, A., Timoneda, N., Bofill-Mas, S., Girones, R. & Rusiñol, M. (2021). NGS Techniques Reveal a High Diversity of RNA Viral Pathogens and Papillomaviruses in Fresh Produce and Irrigation Water. *Foods*, 10 (8), 1820.

<https://doi.org/10.3390/foods10081820>

- **Itarte, M.**, Forés, E., Martínez-Puchol, S., Scheiber, L., Vázquez-Suñé, E., Bofill-Mas, S. & Rusiñol, M. (2024). Exploring viral contamination in urban groundwater and runoff. Article en procés de revisió a la revista *Science of the Total Environment*.

- **Itarte, M.**, Calvo, M., Martínez-Frago, L., Mejías-Molina, C., Martínez-Puchol, S., Girones, R., Medema, G., Bofill-Mas, S. & Rusiñol, M. (2024). Assessing Environmental Exposure to Viruses in Wastewater Treatment Plant and Swine Farm Scenarios with Next-generation Sequencing and Occupational Risk Approaches. *International Journal of Hygiene and Environmental Health*, 259, 114360.

<https://doi.org/10.1016/j.ijheh.2024.114360>

2. Informe de coautoria

- **Itarte, M.**, Bofill-Mas, S., Martínez-Puchol, S., Torrell, H., Ceretó, A., Carrasco, M., Forés, E., Canela, N., Girones, R. & Rusiñol, M. (2021). Looking for a needle in a haystack. SARS-CoV-2 variant characterization in sewage. *Current Opinion in Environmental Science & Health*, 24, 100308. <https://doi.org/10.1016/j.coesh.2021.100308>

L'estudi ha format part del projecte VirWASTE (2020PANDE00044), dins de la convocatòria PANDÈMIES 2020 de l'Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR), i del projecte PROCEED (COI-20201289), finançat pel Centre per al Desenvolupament Tecnològic Industrial (CDTI). La doctoranda ha participat activament en la recerca bibliogràfica, el disseny experimental, la realització dels experiments, i l'anàlisi dels resultats preliminars que s'inclouen en l'article, així com en la seva redacció, amb la supervisió de les seves directores de tesi.

- Martínez-Puchol, S., **Itarte, M.**, Rusiñol, M., Forés, E., Mejías-Molina, C., Andrés, C., Antón, A., Quer, J., Abril, J. F., Girones, R. & Bofill-Mas, S. (2021). Exploring the diversity of coronavirus in sewage during COVID-19 pandemic: Don't miss the forest for the trees. *Science of The Total Environment*, 800, 149562. <https://doi.org/10.1016/j.scitotenv.2021.149562>

L'estudi ha format part del projecte MetaPatFood (AGL2017-86797-C2-1-R) del Ministeri d'Economia, Indústria i Competitivitat (MINECO) i del monitoratge d'aigua residual per a l'estudi de la incidència de SARS-CoV-2 en els que la doctoranda ha participat activament. La doctoranda ha col·laborat en la metodologia i el processament de les mostres, així com en la redacció i revisió de l'article, sota la supervisió de les seves directores de tesi i de la investigadora principal del projecte, la Dra. Rosina Girones. Una versió prèvia de l'article va ser inclosa a la tesi presentada per Sandra Martínez Puchol l'any 2020, amb el títol "SARS-CoV-2 is a new member of urban wastewater virome", en fase de revisió a la revista *Genes* en aquell moment.

- Martínez-Puchol, S., Tarradas-Aleman, M., Mejías-Molina, C., **Itarte, M.**, Rusiñol, M., Baliellas, J., Abasolo, N., Canela, N., Monastiri, A., López-Roig, M., Serra-Cobo, J., Abril, J. F. & Bofill-Mas, S. (2024). Target enrichment metaviromics enables comprehensive surveillance of coronaviruses in environmental and animal samples. Article en procés de revisió a la revista *Heliyon*.

L'estudi ha format part del projecte EPISARS (544/C/2021), finançat a través de les donacions obtingudes durant el programa televisiu "La Marató de TV3" dedicat a la COVID-19, i també del projecte VIRALERT (PID2021-128014OB-I00), finançat pel Ministeri de Ciència, Innovació i Universitats (MCIN/AEI/10.13039/501100011033). La doctoranda ha participat en la metodologia i el processament de les mostres, i ha col·laborat en la redacció i revisió de l'article, sota la supervisió de les seves directores de tesi.

- **Itarte, M.**, Martínez-Puchol, S., Forés, E., Hundesa, A., Timoneda, N., Bofill-Mas, S., Girones, R. & Rusiñol, M. (2021). NGS Techniques Reveal a High Diversity of RNA Viral Pathogens and Papillomaviruses in Fresh Produce and Irrigation Water. *Foods*. 10 (8), 1820.
<https://doi.org/10.3390/foods10081820>

L'estudi també ha format part del projecte MetaPatFood (AGL2017-86797-C2-1-R) del MINECO. La doctoranda ha participat activament en aquest estudi, incloent-hi el disseny experimental, la realització dels experiments, l'anàlisi de les dades i, finalment, en la redacció de l'article, sota la supervisió de les seves directores de tesi i de la investigadora principal del projecte, la Dra. Rosina Girones.

- **Itarte, M.**, Forés, E., Martínez-Puchol, S., Scheiber, L., Vázquez-Suñé, E., Bofill-Mas, S. & Rusiñol, M. (2024). Exploring viral contamination in urban groundwater and runoff. Article en procés de revisió a la revista *Science of the Total Environment*.

L'article ha format part del projecte URBANWAT (PCI2019-103643), finançat pel MCIN, i del projecte UPWATER (101081807), finançat per la convocatòria HORIZON (Horizon-CL6-2022-ZEROPOLLUTION-01). La doctoranda ha participat activament en aquest estudi, incloent-hi el disseny experimental, els mostreigs, el processament de mostres i l'obtenció de dades, l'anàlisi de dades i, finalment, en la redacció de l'article, sota la supervisió de les seves directores de tesi.

- **Itarte, M.**, Calvo, M., Martínez-Frago, L., Mejías-Molina, C., Martínez-Puchol, S., Girones, R., Medema, G., Bofill-Mas, S. & Rusiñol, M. (2024). Assessing Environmental Exposure to Viruses in Wastewater Treatment Plant and Swine Farm Scenarios with Next-generation Sequencing and Occupational Risk Approaches. *International Journal of Hygiene and Environmental Health*. 259, 114360.

<https://doi.org/10.1016/j.ijheh.2024.114360>

L'estudi també ha format part del projecte EPISARS (544/C/2021), finançat a través de les donacions obtingudes durant el programa televisiu "La Marató de TV3" dedicat a la COVID-19, i també del projecte VIRALERT (PID2021-128014OB-I00), finançat pel MCIN (MCIN/AEI/10.13039/501100011033). La doctoranda ha participat activament en aquest estudi, incloent-hi el disseny experimental, els mostreigs, el processament de les mostres, l'obtenció de les dades i l'anàlisi dels resultats. Finalment, ha redactat l'article sota la supervisió de les seves directores de tesi.

Signat,



Dra. Sílvia Bofill Mas

Barcelona, abril 2024



Dra. Marta Rusiñol Arantegui

Barcelona, abril 2024

3. Informe sobre el factor d'impacte dels articles publicats

Els estudis que integren aquesta tesi doctoral han estat publicats o es troben en procés de revisió en revistes científiques rellevants dins de la línia de recerca en la qual ha estat treballant la doctoranda durant els anys de tesi.

L'article "Looking for a needle in a haystack. SARS-CoV-2 variant characterization in sewage" va ser publicat l'any 2021 a la revista *Current Opinion in Environmental Science and Health*, la qual es situava dins del Q1 i del D2 de l'àrea "Environmental Sciences" amb un factor d'impacte de 8.1 segons Journal Citation Reports™.

L'article "Exploring the diversity of coronavirus in sewage during COVID-19 pandemic: Don't miss the forest for the trees" va ser publicat l'any 2021 a la revista *Science of The Total Environment*, la qual es situava dins del Q1 i del D1 de la categoria "Environmental Sciences" amb un factor d'impacte de 10.754 segons Journal Citation Reports™. L'article "Exploring viral contamination in urban groundwater and runoff" es troba actualment en procés de revisió (abril 2024) a la mateixa revista, amb un factor d'impacte actual de 9.8 i continuant dins del Q1 i del D1 de la categoria "Environmental Sciences".

L'article "Target enrichment metaviromics enables comprehensive surveillance of coronaviruses in environmental and animal samples" també es troba actualment en procés de revisió (abril 2024) a la revista *Heliyon*, la qual es situa dins del Q2 i del D4 de la categoria "Multidisciplinary sciences" amb un factor d'impacte de 4.0 segons Journal Citation Reports™.

L'article "NGS Techniques Reveal a High Diversity of RNA Viral Pathogens and Papillomaviruses in Fresh Produce and Irrigation Water" va ser publicat l'any 2021 a la revista *Foods*, la qual es situava dins del

Q1 i del D3 de la categoria “Food Science & Technology” amb un factor d’impacte de 5.561 segons Journal Citation Reports™.

Finalment, l’article “Assessing Environmental Exposure to Viruses in Wastewater Treatment Plant and Swine Farm Scenarios with Next-generation Sequencing and Occupational Risk Approaches” s’ha publicat aquest any 2024 a la revista *International Journal of Hygiene and Environmental Health*, la qual es situa dins del Q1 i del D2 de la categoria “Public, Environmental & Occupational Health” amb un factor d’impacte de 6.0 segons Journal Citation Reports™.

Signat,



Dra. Sílvia Bofill Mas

Barcelona, abril 2024

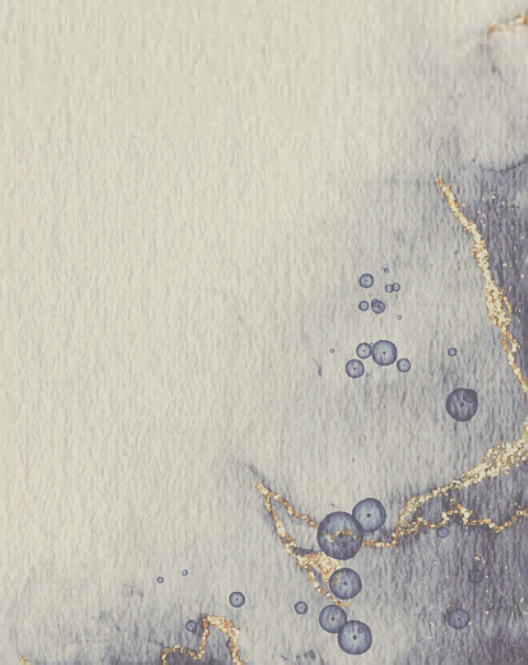


Dra. Marta Rusiñol Arantegui

Barcelona, abril 2024

Capítol 1

L'aigua residual: font de contaminació i eina
per a estudis epidemiològics.
Detecció i caracterització de coronavirus en
un context de pandèmia.



El primer capítol d'aquesta tesi està enfocat a l'estudi de l'aigua residual com a font de contaminació i, a la vegada, com a eina per a la realització d'estudis epidemiològics. En els tres articles inclosos en aquest capítol s'avaluen diferents mètodes de detecció i caracterització dels coronavirus presents en les aigües residuals en un context de pandèmia, especialment mitjançant tècniques de NGS.

En el primer article s'avaluen diferents mètodes per a la detecció i caracterització de variants del SARS-CoV-2 en les aigües residuals, mentre que els altres dos articles es centren en l'aplicació de panells de captura per millorar la detecció i caracterització de coronavirus humans i animals en aigües residuals mitjançant NGS.

Article 1. Buscant una agulla en un paller. Caracterització de variants de SARS-CoV-2 a les aigües residuals

“Looking for a needle in a haystack.

SARS-CoV-2 variant characterization in sewage”

Itarte, M., Bofill-Mas, S., Martínez-Puchol, S., Torrell, H., Ceretó, A., Carrasco, M., Forés, E., Canela, N., Girones, R. i Rusiñol, M.

Current Opinion in Environmental Science & Health. Volume 24, December 2021, 100308; <https://doi.org/10.1016/j.coesh.2021.100308>

Les variants del SARS-CoV-2 han anat emergint arreu del món, i el seu monitoratge és clau per proporcionar alertes primerenques. Aquest treball va consistir en una revisió de les diferents estratègies analítiques utilitzades en mostres d'aigua residual per estudiar la disseminació de les variants del SARS-CoV-2, discutint els avantatges i inconvenients de cadascuna. A més, també es van aportar resultats preliminars de dues aproximacions sensibles i cost-eficaces: assaigs de RT-nPCR específics de variants i una estratègia de seqüenciació massiva d'amplicons de tres regions clau del gen de la proteïna de l'espícula vírica, aquesta última sense ser específica de cap variant concreta.

La revisió i els resultats preliminars presentats van proporcionar informació sobre els avantatges i inconvenients de les diferents estratègies utilitzades per estudiar la disseminació de les variants del SARS-CoV-2 a través de les aigües residuals, incloent-hi aproximacions de RT-qPCR, RT-nPCR i NGS. Les aproximacions basades en PCR i dirigides a mutacions específiques necessiten una actualització constant, però són una opció sensible i cost-eficaç si es vol monitoritzar una variant concreta. En canvi, les tècniques de NGS permeten la detecció simultània de mutacions característiques de diferents variants preocupants (VOC) i

d'interès (VOI) en un sol assaig i són la millor opció per explorar la imatge completa de les variants que circulen en una regió determinada en un moment concret. Les diferents eines utilitzades per identificar i estudiar les variants de SARS-CoV-2 en aigües residuals ofereixen diferents tipus d'informació, la qual cosa és important considerar per seleccionar el mètode més apropiat per a un objectiu establert.

Looking for a needle in a haystack. SARS-CoV-2 variant characterization in sewage

Marta Itarte^{1,2}, Sílvia Bofill-Mas^{1,2}, Sandra Martínez-Puchol^{1,2},
 Helena Torrell³, Adrià Ceretó³, Marina Carrasco¹, Eva Forés^{1,2},
 Núria Canela³, Rosina Girones^{1,2} and Marta Rusiñol⁴

Abstract

SARS-CoV-2 variants are emerging worldwide, and monitoring them is key in providing early warnings. Here, we summarize the different analytical approaches currently used to study the dissemination of SARS-CoV-2 variants in wastewater and discuss their advantages and disadvantages. We also provide preliminary results of two sensitive and cost-effective approaches: variant-specific reverse transcription-nested PCR assays and a nonvariant-specific amplicon deep sequencing strategy that targets three key regions of the viral spike protein. Next-generation sequencing approaches enable the simultaneous detection of signature mutations of different variants of concern in a single assay and may be the best option to explore the real picture at a particular time. Targeted PCR approaches focused on specific signature mutations will need continuous updating but are sensitive and cost-effective.

Addresses

¹ Laboratory of Viruses Contaminants of Water and Food, Genetics, Microbiology & Statistics Department at the University of Barcelona (UB), Barcelona, Catalonia, Spain

² The Water Research Institute (IdRA), Universitat de Barcelona, Barcelona, Catalonia, Spain

³ Eurecat, Centre Tecnològic de Catalunya, Centre for Omic Sciences (COS), Joint Unit Universitat Rovira i Virgili-EURECAT, Unique Scientific and Technical Infrastructures (ICTS), Reus, Spain

⁴ Institute of Environmental Assessment & Water Research (IDAEA), CSIC, Barcelona, Catalonia, Spain

Corresponding author: Rusiñol, Marta (marta.rosinol@idaea.csic.es)

Current Opinion in Environmental Science & Health 2021, 24:100308

This review comes from a themed issue on **COVID-19 in environment: Treatment, Infectivity, Monitoring, Estimation**

Edited by **Manish Kumar, Ryo Honda, Prosun Bhattacharya, Dan Snow and Payal Mazumder**

For a complete overview see the [Issue](#) and the [Editorial](#)

<https://doi.org/10.1016/j.coesh.2021.100308>

2468-5844/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

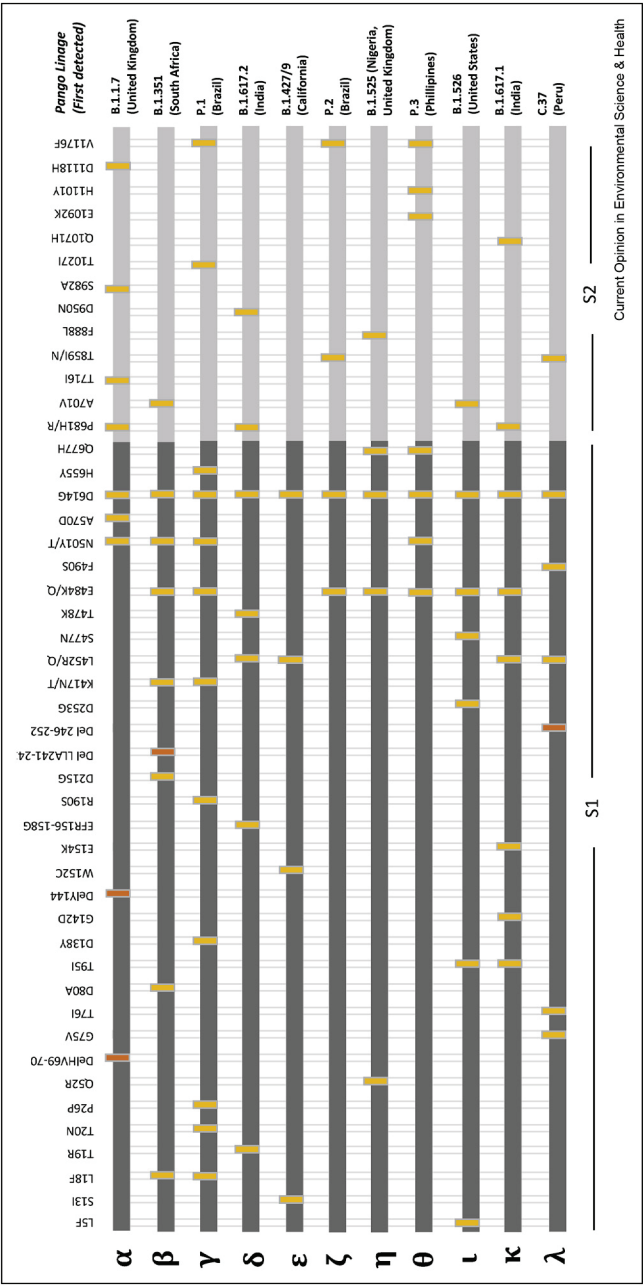
Keywords

SARS-CoV-2, Variants of concern (VOCs), Variants of interest (VOIs), Signature mutations, Wastewater-based epidemiology (WBE), Next-generation sequencing (NGS).

Introduction

Wastewater surveillance for SARS-CoV-2 has proved to be useful in monitoring the evolution of the COVID-19 pandemic. However, new emerging variants are posing new challenges. The SARS-CoV-2 variants α , β , γ and δ (also known as lineages B.1.1.7, B.1.351, P.1 and B.1.617.2, respectively) were first detected in the United Kingdom, South Africa, Brazil and India, respectively, and were immediately considered to be variants of concern (VOCs). Such variants, which have been associated with the fluctuations seen with the pandemic waves, possess mutations that affect viral infectivity and antigenicity. These mutations are mainly located in the gene encoding the viral spike (S) protein. In particular, mutations leading to the E484K and N501Y substitutions within the receptor-binding domain of the S protein have been demonstrated to give the S protein a greater affinity for the human ACE2 receptor [13]. The commonly applied PCR methods used to quantify the concentration of the virus in environmental samples use specific primers and probes targeting the nucleocapsid (N), envelope (E) or RNA-dependent RNA polymerase (RdRp) regions. However, as stated above, the VOCs and the new variants of interest (VOIs) have most of their signature mutations within the S gene. [Figure 1](#) summarizes the signature mutations identified in each VOC and VOI.

Although the combination of genome sequence analysis of samples from COVID-19 patients with epidemiological datasets has produced reliable assessments of the extent of SARS-CoV-2 transmission in the community [22], the time lag between infection and symptoms and the future decrease in sequencing will add further delays compared to the expected immediacy of the results from wastewater surveillance. At the beginning of October 2020, several new SARS-CoV-2 variants started to circulate globally [7]. At that moment, the minimum number of clinical samples that had to be sequenced to find the α variant was 400, assuming that only 5% of the positive clinical samples had been sequenced and that the prevalence of this VOC in the population was 5% [20]. Thus, the analysis of SARS-CoV-2 genomes sequenced from clinical



N501Y signature mutation (present in the α , β , γ and θ variants) in wastewater.

Amplicon sequencing based approaches

Reverse transcription-nested PCR (RT-nPCR) assays followed by Sanger sequencing and/or NGS analysis have been published for SARS-CoV-2 characterization. In October 2020, Martin and collaborators designed an RT-nPCR approach followed by Sanger sequencing and NGS analysis of the amplified products from five different regions of the viral genome, which demonstrated changes in the predominance of the virus variants [20]. La Rosa and coworkers [25] adopted a similar approach involving conventional Sanger sequencing of the amplicon but focusing only on key mutations of the S gene, which allowed rapid screening of the SARS-CoV-2 variants [La Rosa et al. 2021]. Recently, another group from the United Kingdom used two different RT-nPCR assays targeting the RdRP and ORF8b gene regions for diagnostics and two primer sets targeting the S gene regions to discriminate between the α , β and γ variants [28].

Sequencing amplicons using NGS, commonly known as amplicon deep sequencing (ADS), has not only been applied to selected parts of the SARS-CoV-2 genome but also to the whole genome as an informative method for detecting and identifying SARS-CoV-2 variants. Several custom enrichment strategies based on designing primer sets coupled with Illumina-compatible library preparation kits have been used to sequence amplified fragments spanning the whole or near-complete genome of SARS-CoV-2 from environmental samples [2,15,18,20,28]. Other studies have used the open-source ARTIC protocol [3,16,23]. This protocol, released in March 2020 and designed to sequence the virus from clinical samples, uses 98 multiplexing PCR primer pairs to amplify the whole genome of the virus [24]. Similarly, the commercial AmpliSeq SARS-CoV-2 Research Panel (Thermo Fisher Scientific) consists of two pools with amplicons ranging from 125 bp to 275 bp that covers >99% of the SARS-CoV-2 genome and are compatible with either Illumina or Ion Torrent sequencing platforms [2]. Another strategy based on NGS is the use of a commercial oligo-capture approach, like the Illumina Respiratory Virus Oligo Panel (Illumina, Inc.) or the VirCapSeq Enrichment Kit (Roche), which are designed to enrich the sequences of human respiratory or vertebrate viruses, respectively, and both have been applied to complex environmental samples prior to massive sequencing [8,21].

Based on the findings of available studies, the most abundant single nucleotide variations (SNVs) that have been identified in wastewater to date correspond to the most abundant SNVs in clinical samples [8]. The

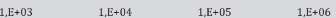
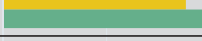
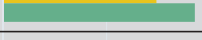
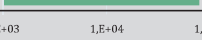
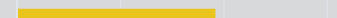
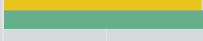
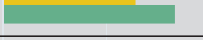
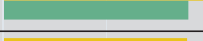
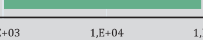
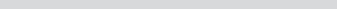
identification of an individual or several signature mutations (Figure 1) located in close proximity to one another within the sample amplicon can help identify new SNVs in the population being analyzed. When using these approaches in environmental samples containing a mixture of variant sequences, there is a possibility of generating artificial genome reconstructions or artefacts during sequence assembly, which could result in unreliable VOC or VOI assignments.

The ADS of selected regions provides a more robust characterization of genomic variants compared to broader genome reconstructions within individual samples. When applied to clinical samples, long-read sequencing platforms have been proven to be efficient in obtaining highly accurate consensus-level sequences despite the higher error rates [4]. However, to our knowledge, this approach has not been applied in the study of SARS-CoV-2 variants in sewage.

Specific regions for the characterization of SARS-CoV-2 genomic variants

Approaches targeting selected regions of the SARS-CoV-2 genome in which signature mutations are located generate more interest compared to the sequencing of other regions that are more conserved and less informative about genomic variants. For discriminating between variants, European authorities have established that sequencing should cover at least the S gene, particularly that encoding the entire N-terminal region and the receptor-binding domain (RBD) corresponding to amino acids 1 to 541 [9]. Preliminary data obtained from two different approaches that were developed by our research group are detailed below. These approaches involved specific RT-nPCR assays targeting the signature mutations of the main VOCs and VOIs followed by Sanger sequencing (assay A and B) and an ADS strategy targeting three different regions of the S gene (assays A1, A2 and A3). Both approaches were tested in parallel in samples collected from February to May 2021 from wastewater treatment plants (WWTP) of different sizes located in Catalonia, northeast Spain. More information about the methodology is provided in the Supplementary Material. The results obtained are summarized in Table 1 and the datasets generated are available in Zenodo under the DOI number <https://doi.org/10.5281/zenodo.5497909>.

As indicated in Table 1, Sanger sequencing allowed the identification of signature mutations in the samples, in which the following were predominant: Del69/70, Del144, K417N and E484K. The ADS approach gave information about the genomic diversity in each sample, showing different signature mutation combinations that are compatible with different variants as expected in mixtures coming from wastewater samples. Interestingly, ADS indicated the moment when the α variant

Table 1					
Summary of SARS-CoV-2 concentrations (GC/L) detected using RT-qPCR and signature mutations detected using RT-nPCR and Sanger sequencing or ADS in a MiSeq platform. ND: not detected.					
			RT-nPCR	ADS	
WWTP 1 (1,497,767 inh.)	2/2/2021		K417N E484K	Del69/70 (0.1%) S477N (14.6%) D614G (73.3%) N501Y A570D D614G (26.7%)	
	9/2/2021		Del69/70 Del144 E484K	Del69/70 (100%) N501Y A570D D614G (100%)	
	13/4/2021		Del69/70 Del144 K417N E484K	Del69/70 (100%)	
	20/4/2021		K417N E484K	Del69/70 (99.8%)	
	4/5/2021		Del69/70 Del144	Del69/70 (100%)	
					
WWTP 2 (183,517 inh.)	2/2/2021		ND	Del69/70 (64.1%) D614G (44.6%) N501Y A570D D614G (55.4%)	
	9/2/2021		Del69/70 Del144 E484K	Del69/70 (99.5%) N501Y A570D D614G (100%)	
	13/4/2021		E484K	Del69/70 (99.5%)	
	20/4/2021		ND	Del69/70 (99.5%)	
	4/5/2021		ND	ND	
					
WWTP 3 (68,860 inh.)	2/2/2021		ND	D614G (100%)	
	9/2/2021		ND	Del69/70 (100%) D614G (100%)	
	13/4/2021		Del69/70 Del144 E484K	Del69/70 (99.6%)	
	20/4/2021		ND	ND	
	4/5/2021		ND	Del69/70 (100%)	
					
					

probably became predominant in Catalonia. From all the sequences obtained from the NGS analysis of the samples collected on 2nd February 2021, the Del69/70 mutation was present in 0.1%, 64.1% and 0% of the sequences obtained from WWTP1, WWTP2 and WWTP3, respectively. One week later, these percentages increased to 99.5%–100%, which was also observed in other signature mutations of the α variant (N501Y, A570D and D614G). These ADS results associated with

α variant predominance agree with the information obtained from the Sanger sequencing. The detection of the signature mutations compatible with the α variant with Sanger sequencing was only possible in samples that showed a high percentage of the signature mutations of the α variant by ADS, or in other words, when these mutations were predominant among the mixture of sequences. The other signature mutation identified by ADS was S477N, which is characteristic of the ι variant.

Table 2

List of pros and cons of the different methodologies used in the study of SARS-CoV-2 variants in sewage samples.

Method	Pros	Cons	Applicability
RT-qPCR	Low cost Approximation of the specific signature mutation vs. WT proportion in a mixture Fast obtention of results Detects rare mutations and discriminates closely related sequences	Different target sensitivities when multiplexing Designed to detect a signature mutation of a specific variant only, thus not giving information about other possible variants also present in the sample	Monitoring of a specific variant in a region where it has spread
RT-ddPCR	Fast obtention of results Approximation of the specific signature mutation vs. WT proportion in a mixture Detects rare mutations and discriminates closely related sequences More sensitive than RT-qPCR	Designed to detect a signature mutation of a specific variant only, thus not giving information about other possible variants also present in the sample More expensive than RT-qPCR	More sensitive monitoring of a specific variant in a region where it has spread
RT-nPCR + Sanger Sequencing	Low cost Fast obtention of results Easy interpretation of the results May use primers specific targeting defined signature mutations	Detecting only the predominant variant in the mixture Not quantitative Cannot be effectively performed in conditions of low virus titers	Fast elucidation of the predominant variant circulating in a region
NGS	Shows the diversity of variants circulating More extensive information about mutations in a larger range of the genome	Expensive Extensive bioinformatics analysis Not quantitative Labour intensive Time consuming Might lead to artificial consensus genomes Cannot be effectively performed in conditions of low virus titers	Characterization of variant diversity circulating in a region

Variant study approaches: the pros and cons

Different analytical approaches for the study of SARS-CoV-2 variants in wastewater samples have been developed, each one providing different types of information. In Table 2, the pros and cons of the different methodologies that have been used to date are listed. Depending on their intrinsic properties, a suitable application has been suggested.

RT-qPCR and RT-ddPCR are designed to detect a signature mutation of a particular variant and are the fastest at providing results. Both methodologies are often designed as duplex or multiplex, allowing the simultaneous detection of other variants and giving an estimation of their percentages among other simultaneously occurring variants. Thus, they are appropriate for monitoring a specific variant in a region where it has spread and become established since a certain proportion of the target variant with respect to the others is needed to be detected. RT-ddPCR might be more sensitive and precise than RT-qPCR, but it is also more expensive [1, 6,14] [Abachin_et_al_2017].

However, wastewater is a complex sample, and it is likely to contain a mixture of variants. In a region where the predominant variant circulating within the population is not clear or where the situation is constantly

changing, non-variant-specific methodologies might be more suitable since they do not need continuous updating of the assay. In such cases, RT-nPCR assays followed by Sanger sequencing of specific regions containing signature mutations would be highly informative and would identify the predominant variant circulating in the population, as this type of sequencing gives information about the most abundant sequence amplified. Furthermore, RT-nPCR can use specific primers for a defined mutation that can target specific variants and regions where other mutations may occur. By contrast, if the objective is to perform an accurate characterization of the diversity present in wastewater, or in other words, identify different variants present in a mixture, NGS analysis would be more appropriate. The extensive information provided by NGS techniques, considered to be expensive, requires an exhaustive bioinformatics analysis and expertise.

Conclusions

Monitoring SARS-CoV-2 variants in wastewater is important for epidemiological surveillance in a community. Different analytical approaches have been developed to identify and study the dissemination of SARS-CoV-2 variants in wastewater samples, including RT-qPCR, RT-nPCR, and NGS approaches. Due to their intrinsic nature, each method has pros and cons and provides different types of information that is

important to consider when selecting the appropriate method for a specific objective. In a postpandemic scenario, when PCR-based assays and sequencing of clinical samples will decrease, the sequencing of a subset of wastewater samples may be enough to monitor the circulation of different VOCs and VOIs in a community. A representative sample needs to be collected regularly from a certain region to accurately estimate and monitor the prevalence of SARS-CoV-2 variants. Nonvariant-specific techniques may be the best option to explore the real picture of all the circulating variants at a particular time, providing broader information that can contribute to community surveillance. This study provides guidance on available approaches for detecting and identifying circulating SARS-CoV-2 variants considering different scenarios. Further work on the application of massive sequencing of SARS-CoV-2 from environmental samples is needed towards producing longer fragments in order to avoid overlapping and chimera constructions, and also shorter bioinformatic processing for an effective early warning.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was partially supported by the Catalan Government through 2020PANDE00044, and CDTI project PROCEED, EXP - 00132149/COI-20201289. Sílvia Bofill-Mas is a Serra-Hunter fellow at the University of Barcelona. Eva Forés is an APIF fellow at the University of Barcelona. Marta Itarte is a fellow of the Catalan Government 'AGAUR' (FI) at the University of Barcelona. We also gratefully acknowledge the authors, originating and submitting laboratories of the SARS-CoV-2 sequences from GISAID's EpiCov Database (<https://www.epicov.org/epi3/>) used in the phylogenetic analyses; access to the individual isolates is facilitated through the GISAID website (<https://www.gisaid.org/>). Graphical Abstract was created with [Biorender.com](https://www.biorender.com).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.coesh.2021.100308>.

References

Papers of particular interest, published within the period of review, have been highlighted as:

- * of special interest
 - ** of outstanding interest
1. Abachin E, Convers S, Falque S, Esson R, Mallet L, Nougarede N: **Comparison of reverse-transcriptase qPCR and droplet digital PCR for the quantification of dengue virus nucleic acid.** *Biologicals* 2018, **52**:49–54, <https://doi.org/10.1016/j.biologicals.2018.01.001>.
 2. Agrawal S, Orschler L, Schubert S, Zachmann K, Heijnen L: **A pan-European study of SARS-CoV-2 variants in wastewater under the EU Sewage Sentinel System.** *medRxiv Prepr* 2021, **25**.
 3. Bar-or I, Weil M, Indenbaum V, Bucris E, Bar-ilan D, Elul M, Levi N, Aguvaev I, Cohen Z, Shirazi R, Erster O, Sela-brown A, Sofer D, Mor O, Mendelson E, Zuckerman NS: **Detection of SARS-CoV-2 variants by genomic analysis of wastewater samples in Israel.** *Sci Total Environ* 2021, **789**:148002, <https://doi.org/10.1016/j.scitotenv.2021.148002>.
 4. Bull RA, Adikari TN, Ferguson JM, Hammond JM, Stevanovski I, Beukers AG, Naing Z, Yeang M, Verich A, Gamaarachchi H, Kim KW, Luciani F, Stelzer-Braid S, Eden JS, Rawlinson WD, van Hal SJ, Deveson IW: **Analytical validity of nanopore sequencing for rapid SARS-CoV-2 genome analysis.** *Nat Commun* 2020, **11**:1–8, <https://doi.org/10.1038/s41467-020-20075-6>.
 5. Carcereny A, Martínez-velázquez A, Bosch A, Allende A, Truchado P: **Monitoring emergence of SARS-CoV-2 B.1.1.7 variant through the Spanish national SARS-CoV-2 wastewater surveillance system (VATar COVID-19) from december 2020 to March.** *medRxiv Prepr* 2021.
 6. Ciesielski M, Blackwood D, Clerkin T, Gonzalez R, Thompson H, Larson A, Noble R: **Assessing sensitivity and reproducibility of RT-ddPCR and RT-qPCR for the quantification of SARS-CoV-2 in wastewater.** *J Virol Methods* 2021, **297**:114230, <https://doi.org/10.1016/j.jviromet.2021.114230>.
 7. CDC: **Science brief : emerging SARS-CoV-2 variants.** 2021.
 8. Crits-Christoph A, Kantor RS, Olm MR, Whitney ON, Al-Shayeb B, Lou YC, Flamholz A, Kennedy LC, Greenwald H, Hinkle A, Hetzel J, Spitzer S, Koble J, Tan A, Hyde F, Schroth G, Kuersten S, Banfield JF, Nelson KL: **Genome sequencing of sewage detects regionally prevalent SARS-CoV-2 variants.** *mBio* 2021, **9**.
 9. ECDC and WHO: **Methods for the detection and identification of SARS-CoV-2 variants.** *Stockholm and Copenhagen.* 2021.
 10. European Commission: **HERA Incubator: anticipating together the threat of COVID-19 variants.** 2021.
 11. Gawlik B, Tavazzi S, Mariani G, Skejo H, Sponar N, Higgins T, Medema G, T., W.: **SARS-CoV-2 Surveillance employing Sewage Towards a Sentinel System.** 2021, <https://doi.org/10.2760/300580>.
- Summarises methodologies and entailed costs for a European Sewage Sentinel System for SARS-CoV-2.
12. Graber TE, Mercier É, D'aoust PM, Hoang H-D, Tian X, Taaneem S, Bhatnagar K, Delatolla R: **An allele-specific primer extension assay to quantify the proportion of B.1.1.7-specific SARS-CoV-2 RNA in wastewater.** *medRxiv* 2021, **2021.02.22.21252041**.
 13. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, Ludden C, Reeve R, Rambaut A, Peacock SJ, Robertson DL: **SARS-CoV-2 variants, spike mutations and immune escape.** *Nat Rev Microbiol* 2021, **19**:409–424, <https://doi.org/10.1038/s41579-021-00573-0>.
 14. Heijnen L, Elsinga G, Graaf de M De, Molenkamp R, Koopmans MPG, Medema G: **Droplet Digital RT-PCR to detect SARS-CoV-2 variants of concern in wastewater.** *medRxiv Prepr* 2021, **14**.
 15. Izquierdo-Lara R, Elsinga G, Heijnen L, Oude Munnink BB, Schapendonk CME, Nieuwenhuijsen D, Kon M, Lu L, Aarestrup FM, Lycett S, Medema G, Koopmans MPG, De Graaf M: **Monitoring SARS-CoV-2 circulation and diversity through community wastewater sequencing, The Netherlands and Belgium.** *Emerg Infect Dis* 2021, **27**:1405–1415, <https://doi.org/10.3201/eid2705.204410>.
 16. Jahn K, Dreifuss D, Topolsky I, Kull A, Stachler E, Stadler T, Ort C, Kohn T, Julian TR, Beerenwinkel N: **Detection of SARS-CoV-2 variants in Switzerland by genomic analysis of wastewater samples.** 2021:1–14.
 17. Kazama S, Miura T, Masago Y, Konta Y, Tohma K, Manaka T, Liu X, Nakayama D, Tanno T, Saito M, Oshitani H, Omura T: **Environmental surveillance of norovirus genogroups I and II for sensitive detection of epidemic variants.** *Appl Environ Microbiol* 2017, **83**, <https://doi.org/10.1128/AEM.03406-16>.
 18. Ko K, Nagashima S, Bunthen E, Ouoba S, Akita T, Sugiyama A, Ohisa M, Sakaguchi T, Tahara H, Ohge H, Ohdan H, Kubo T, Kishita E, Kuwabara M, Takahashi K, Tanaka J: **Molecular characterization and the mutation pattern of SARS-CoV-2 during first and second wave outbreaks in Hiroshima, Japan.**

PLoS One 2021, **16**:1–15, <https://doi.org/10.1371/journal.pone.0246383>.

19. Lee WL, McElroy KA, Armas F, Imakaev M, Gu X, Duvallet C, Chandra F, Chen H, Leifels M, Xiao A, Moniz K, Ghaeli N, Matus M: **Quantitative detection of SARS-CoV-2 B.1.1.7 variant in wastewater by allele-specific RT- qPCR.** *medRxiv Prepr* 2021:1–32.
20. Martin J, Klapsa D, Wilton T, Zambon M, Bentley E, Bujaki E, Fritzsche M, Mate R, Majumdar M: **Tracking SARS-CoV-2 in sewage : evidence of changes in virus variant predominance during COVID-19 pandemic.** *Viruses* 2020, **12**.
 This study described a novel RT-nested PCR that allowed tracking changes in virus variant predominance.
21. Martínez-Puchol S, Itarte M, Rusiñol M, Forés E, Mejías-Molina C, Andrés C, Antón A, Quer J, Abril JF, 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, <https://doi.org/10.1016/j.scitotenv.2021.149562>.
22. Oude Munnink BB, Nieuwenhuijse DF, Stein M, O'Toole Á, Haverkate M, Mollers M, Kamga SK, Schapendonk C, Pronk M, Lexmond P, van der Linden A, Bestebroer T, Chestakova I, Overmars RJ, van Nieuwkoop S, Molenkamp R, van der Eijk AA, GeurtsvanKessel C, Vennema H, Meijer A, Rambaut A, van Dissel J, Sikkema RS, Timen A, Koopmans M, Oudehais GJAPM, Schinkel J, Kluytmans J, Kluytmans-van den Bergh M, van den Bijllaardt W, Bertelvelsen RG, van Rijen MML, Schneeberger P, Pas S, Diederien BM, Bergmans AMC, van der Eijk PAV, Verweij J, Buiting AGN, Streefkerk R, Aldenkamp AP, de Man P, Koelema JGM, Ong D, Paltansing S, Veassen N, Steven J, Bakker L, Brockhoff H, Rietveld A, Slijkerman Megelink F, Cohen Stuart J, de Vries A, van der Reijden W, Ros A, Lodder E, Verspui-van der Eijk E, Huijskens I, Kraan EM, van der Linden MPM, Debast SB, Naiemi AI N, Kroes ACM, Damen M, Dinant S, Lekkerkerk S, Pontesilli O, Smit P, van Tienen C, Godschalk PCR, van Pelt J, Ott A, van der Weijden C, Wertheim H, Rahamat-Langendoen J, Reimerink J, Bodewes R, Duizer E, van der Veer B, Reusken C, Lutgens S, Schneeberger P, Hermans M, Wever P, Leenders A, ter Waarbeek H, Hoebe C: **Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in The Netherlands.** *Nat Med* 2020, **26**:1405–1410, <https://doi.org/10.1038/s41591-020-0997-y>.
23. Pérez-Cataluña A, Chiner-Oms Á, Cuevas-Ferrando E, Díaz-Reolid A, Falcó I, Randazzo W, Girón-Guzmán I, Allende A, Bracho MA, Comas I, Sánchez G: **Detection of genomic variants of SARS-CoV-2 circulating in wastewater by high-throughput sequencing.** *medRxiv Prepr* 2021.
24. Quick J: **nCoV-2019 sequencing protocol v3 (LoCost).** *protocols.io* 2020, **3**:1–24.
25. La Rosa I, Mancini P, Ferraro GB, Veneri C, Iaconelli M, Lucentini L, Bonadonna L, Brusaferrero S, Brandtner D, Fasanella A, Pace L, Parisi A, Galante D, Suffredini E: **Rapid screening for SARS-CoV-2 variants of concern in clinical and environmental samples using nested RT-PCR assays targeting key mutations of the spike protein.** *Water Res* 2021, **197**:117104, <https://doi.org/10.1016/j.watres.2021.117104>.
 This study described an RT-nested PCR assay for rapid screening for SARS-CoV-2 variants targeting key mutations of the spike protein by Sanger sequencing.
26. Vogels CBF, Breban MI, Ott IM, Id TA, Petrone E, Watkins AE, Id CCK, Id RE, Id ER, Goes J, Id DJ, Claro IM, Magalh G, Id F, Crispim MAE, Network BCG, Id LS, Tegally H, Anyaneji UJ, Genomic N, Africa S, Hodcroft EB, Id CEM, Khullar G, Metti J, Id JTD, Mackay MJ, Nash M, Wang J, Liu C, Hui P, Murphy S, Neal C, Laszlo E, Landry L, Muyombwe A, Downing R, Razeq J, Oliveira de T De: **Multiplex qPCR discriminates variants of concern to enhance global surveillance of.** *PLoS Biol* 2021, **351**:1–12, <https://doi.org/10.1371/journal.pbio.3001236>.
 The study described the first multiplex RT-qPCR published that discriminates variants of concern to enhance global surveillance of SARS-CoV-2.
27. Wang H, Miller JA, Verghese M, Sibai M, Solis D, Mfuh KO, Jiang B, Iwai N, Mar M, Huang C, Yamamoto F, Sahoo MK, Zehnder J, Pinsky BA: **Multiplex SARS-CoV-2 genotyping RT-PCR for population-level variant screening and epidemiologic surveillance.** *J Clin Microbiol* 2021, **50**, <https://doi.org/10.1128/JCM.00859-21>.
28. Wilton T, Bujaki E, Klapsa D, Majumdar M, Zambon M, Fritzsche M, Mate R, Martin J: **Rapid increase of SARS-CoV-2 variant B.1.1.7 detected in sewage samples from england between october 2020 and january 2021.** *Am. Soc. Microbiol.* 2021, **6**, <https://doi.org/10.23959/sfddtj-1000004>.
29. Wurtzer S, Waldman P, Levert M, Mouchel JM, Gorgé O, Boni M, Maday Y, Marechal V, Moulin L: **Monitoring the propagation of SARS CoV2 variants by tracking identified mutation in wastewater using specific RT-qPCR.** *medRxiv* 2021. 2021.03.10.21253291.

Article 2. Explorant la diversitat de coronavirus en les aigües residuals durant la pandèmia de la COVID-19: que els arbres no ens impedeixin veure el bosc

“Exploring the diversity of coronavirus in sewage during COVID-19 pandemic: Don't miss the forest for the trees”

Martínez-Puchol, S., **Itarte, M.**, Rusiñol, M., Forés, E., Mejías-Molina, C., Andrés, C., Antón, A., Quer, J., Abril, J. F., Girones, R. i Bofill-Mas, S.

Science of The Total Environment. Volume 800, 15 December 2021, 149562; <https://doi.org/10.1016/j.scitotenv.2021.149562>

Amb l'arribada de la pandèmia de la COVID-19, les tècniques de NGS han demostrat ser una eina important per a la caracterització genètica del SARS-CoV-2 a partir de mostres clíniques. L'ús de diferents tècniques de NGS disponibles aplicades a mostres d'aigua residual podria ser la clau per estudiar amb profunditat el viroma excretat, no només centrant-se en la circulació i la tipificació del SARS-CoV-2, sinó també per a la detecció d'altres virus potencialment pandèmics dins de la mateixa família *Coronaviridae*.

En aquest estudi, es van seqüenciar mostres compostes de 24 hores d'aigua residual, obtingudes durant els mesos de març i juliol de 2020, mitjançant dues tècniques de NGS, amb i sense enriquiment de dianes. La seqüenciació massiva amb enriquiment de dianes es va realitzar mitjançant la captura de seqüències de virus de vertebrats amb el panell VirCapSeq-VERT.

La seqüenciació massiva amb el panell de captura VirCapSeq-VERT va permetre obtenir el viroma de les mostres d'aigua residual analitzades i detectar virus humans de la família *Coronaviridae*, incloent-hi SARS-CoV-2 i HCoV-OC43. A més, es van detectar altres coronavirus animals amb aquesta tècnica, concretament coronavirus que infecten animals domèstics i rosegadors.

Els resultats obtinguts reforcen la idea que el SARS-CoV-2 és un membre més del viroma de l'aigua residual i, a més, demostren que existeix una cocirculació de coronavirus humans i animals. La metodologia basada en l'ús d'un panell de captura dirigit a una àmplia varietat de famílies víriques sembla una bona opció per avaluar la cocirculació de coronavirus humans i animals, tant coneguts com desconeguts, la qual cosa podria ser rellevant pel que fa a zoonosis potencials i el descobriment de virus nous. Per altra banda, els resultats obtinguts suggereixen que panells de captura més específics, dirigits a la captura de SARS-CoV-2 o altres coronavirus, serien més apropiats per a estudiar la diversitat genètica de les variants del SARS-CoV-2 i d'altres coronavirus circulants.



Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Exploring the diversity of coronavirus in sewage during COVID-19 pandemic: Don't miss the forest for the trees

Sandra Martínez-Puchol^{a,b,*}, Marta Itarte^{a,b}, Marta Rusiñol^c, Eva Forés^{a,b}, Cristina Mejías-Molina^a, Cristina Andrés^d, Andrés Antón^d, Josep Quer^e, Josep F. Abril^f, Rosina Girones^{a,b}, Sílvia Bofill-Mas^{a,b}

^a Laboratory of Viruses Contaminants of Water and Food, Genetics, Microbiology & Statistics Dept., Universitat de Barcelona, Barcelona, Catalonia, Spain

^b The Water Research Institute (IdRA), Universitat de Barcelona, Barcelona, Catalonia, Spain

^c Institute of Environmental Assessment & Water Research (IDAEA), CSIC, Barcelona, Catalonia, Spain

^d Respiratory Viruses Unit, Virology Section, Microbiology Department, Hospital Universitari Vall d'Hebron, Vall d'Hebron Research Institute, Universitat Autònoma de Barcelona, Barcelona, Catalonia, Spain

^e Liver Unit, Liver Diseases - Viral Hepatitis, Vall d'Hebron Institut de Recerca (VHIR), Vall d'Hebron Hospital Universitari, Vall d'Hebron Barcelona Hospital Campus, Passeig Vall d'Hebron 119-129, 08035 Barcelona, Spain

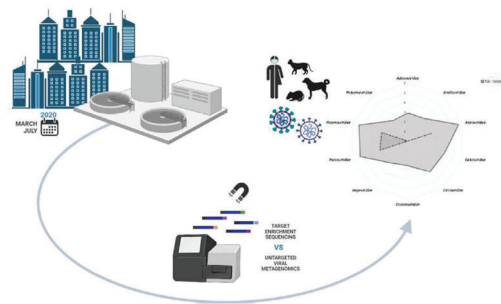
^f Computational Genomics Lab, Genetics, Microbiology & Statistics Dept., Universitat de Barcelona; Institut de Biomedicina (IBUB), Barcelona, Catalonia, Spain



HIGHLIGHTS

- NGS methods have been applied to study the sewage virome during COVID-19 pandemic.
- *Coronaviridae* sequences were not detected in sewage using untargeted metagenomics.
- Target Enrichment provided with SARS-CoV-2 sequences as part of the sewage virome.
- Human and animal CoV co-circulation in sewage only detected with Target Enrichment.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 9 June 2021

Received in revised form 5 August 2021

Accepted 6 August 2021

Available online 9 August 2021

Keywords:

Next generation sequencing

Viral metagenomics

Coronavirus

Sewage virome

Target enrichment sequencing

ABSTRACT

In the wake of the COVID-19 pandemic, the use of next generation sequencing (NGS) has proved to be an important tool for the genetic characterization of SARS-CoV-2 from clinical samples. The use of different available NGS tools applied to wastewater samples could be the key for an in-depth study of the excreted virome, not only focusing on SARS-CoV-2 circulation and typing, but also to detect other potentially pandemic viruses within the same family. With this aim, 24-hours composite wastewater samples from March and July 2020 were sequenced by applying specific viral NGS as well as target enrichment NGS. The full virome of the analyzed samples was obtained, with human *Coronaviridae* members (CoV) present in one of those samples after applying the enrichment. One contig was identified as HCoV-OC43 and 8 contigs as SARS-CoV-2. CoVs from other animal hosts were also detected when applying this technique. These contigs were compared with those obtained from contemporary clinical specimens by applying the same target enrichment approach. The results showed that there is a co-circulation in urban areas of human and animal coronaviruses infecting domestic animals and rodents. NGS enrichment-based protocols might be crucial to describe the occurrence and genetic characteristics of SARS-CoV-2 and other *Coronaviridae* family members within the excreted virome present in wastewater.

© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author at: Laboratory of Viruses Contaminants of Water and Food, Genetics, Microbiology & Statistics Dept., Universitat de Barcelona, Barcelona, Catalonia, Spain.
E-mail address: smartinezpuchol@ub.edu (S. Martínez-Puchol).

1. Introduction

SARS-CoV-2 was identified in China at the end of 2019 and has become the first pandemic coronavirus (CoV) (Wu et al., 2020a, 2020b). Upon confirmation that COVID-19 patients shed SARS-CoV-2 in feces, different studies provided significant correlation between the concentration of SARS-CoV-2 in wastewater and the prevalence of COVID-19 in the served population, increasing the evidence that wastewater is a good indicator of the prevalence of the excreted virus in a population (wastewater-based epidemiology, WBE). SARS-CoV-2 RNA has been reported in feces of 55% of COVID-19 patients in the study cohort, prolonged fecal shedding was also reported with fecal samples remaining positive for nearly 5 weeks after the patients' respiratory samples tested negative for SARS-CoV-2 RNA (Wu et al., 2020a, 2020b). The concentration of SARS-CoV-2 excreted from infected patients has been estimated to be in the range of $1\text{E}+02$ to $1\text{E}+07$ copies per gram of feces (Wölfel et al., 2020).

It is then presumable that, in the context of a pandemic, SARS-CoV-2 could become a member of wastewater virome where viruses commonly present in human excreta have been described to occur (Cantalupo et al., 2011; Ng et al., 2012). The presence of this virus in wastewater samples has been extensively reported by RT-PCR based methods. Also, an extensive effort for SARS-CoV-2 sequencing has been done based on metagenomic approaches focused on variant analysis in both clinical and sewage samples. However, the use of metagenomic approaches to describe its occurrence together with other excreted viruses in the context of sewage virome has not been described. Knowing that SARS-CoV-2 is found in sewage in moderate concentrations and sewage is a complex matrix comprising a wide variety of viruses, studying SARS-CoV-2 in this context may pose a challenge. Previous studies report procedures to describe wastewater virome and compare high throughput viral metagenomics to target enrichment and amplicon deep sequencing approaches (Martínez-Puchol et al., 2020). In this study, we applied a probe-capture target enrichment NGS for describing wastewater virome from 2 wastewater 24-h composite samples collected in March and July 2020 from a wastewater treatment plant located in the city of Barcelona, when COVID-19 incidence was of approximately 258 cases and 175 cases/100,000 inhabitants respectively. The results obtained were compared with those obtained from clinical samples collected at the same time and sequenced applying the same enrichment approach.

2. Materials and methods

2.1. Sample collection and viral concentration

Urban 24-h composite sewage samples were collected on March 19th and July 14th 2020 from a wastewater treatment plant (WWTP) located in the city of Barcelona that serves up to 2.8 million population equivalents and receives domestic and industrial waste from the sewer system. Samples were collected and kept at 4 °C in a sterile container until viral particles from 70 ml of sewage were concentrated by centrifugal ultrafiltration. After a debris removal (15 min, 4500 ×g), the samples were ultrafiltered with a Centricon® Plus-70 device (30 kDa) following the manufacturer instructions, obtaining a viral concentrates of 200 µl. Additionally, two SARS-CoV-2-positive naso/oropharyngeal swabs were obtained on March 15th from a male (clinical sample A) and on March 24th from a female (clinical sample B) patients attended at Hospital Universitari Vall d'Hebron (HUVH).

2.2. SARS-CoV-2 quantification by (RT)-qPCR

Nucleic acid extraction (NA) was performed as described previously (Fernandez-Cassi et al., 2018) with QIAamp Viral RNA Mini Kit. The concentration of SARS-CoV-2 RNA in wastewater samples was measured by (RT)-qPCR of two viral targets in nucleocapsid phosphoprotein (N1 and

N2 region) and clinical samples in HUVH by commercial real-time multiplex RT-PCR (Allplex™ 2019-nCoV Assay, Seegene, South Korea). EURM-019 single stranded RNA (ssRNA) fragments of SARS-CoV-2 (Joint Research Centre, EC, <https://crm.jrc.ec.europa.eu/p/EURM-019>) were used to construct the standard for quantitation. JC polyomavirus (JCPyV) was quantified in the samples as an indicator of human fecal viral contamination as previously described (Bofill-Mas et al., 2006).

2.3. Library construction and probe-based capture of viral sequences

Before library preparation, NA were retrotranscribed to cDNA, tagged and complemented to obtain dsDNA. This viral randomly tagged dsDNA was then amplified (25 cycles) to obtain the sufficient amount of DNA for the preparation of libraries, as previously described (Fernandez-Cassi et al., 2018). Libraries were prepared in duplicate using KAPA HyperPrep Kit following the instructions provided by the manufacturer (Roche-Kapa Biosystems). One replicate of the libraries was hybridized with probes designed to capture sequences from vertebrate viral pathogens (VirCapSeq Enrichment Kit, Roche). Two negative controls were also processed, one of them with the enrichment kit.

After the capture, quality and concentration were re-checked and sequencing of the libraries from the captured, non-captured, and negative controls was performed (Illumina Miseq 2x300bp).

Clinical samples were sequenced in an independent Illumina Miseq 2x300bp run (in this case, consensus sequences were assembled with reads produced from another Illumina TruSeq high coverage sequencing run on same samples, and are already available at GISAID as EPI_ISL_418860 and EPI_ISL_418861, respectively).

2.4. Bioinformatic analysis

The sequencing raw data obtained was analyzed with Genome Detective Virus Tool (Vilsker et al., 2018) and the contigs with a nucleotide identity $\geq 70\%$ (when comparing against the known viral genomes database) were further processed with Geneious (v11.1.5; <https://www.geneious.com>). Simultaneously, raw reads from wastewater and clinical samples were cleaned of technical sequences and trimmed by quality using Trimmomatic (v0.38; (Bolger et al., 2014)) in order to remove low quality segments and Illumina adapters (min Phred score = Q20 on 4 bp window, min read length = 30 bp, leading/trailing clip = 15 bp, max mismatch count = 2, palindrome clip threshold = 30, and simple clip threshold = 10). When one of the reads, either R1 or R2, was discarded, the remaining one was collected into a single-ended (SG) reads file to use along with the resulting filtered paired-end reads (PE) reads later on. Samtools (v1.9; [11]) and bamtools (v2.5.1; (Barnett et al., 2011)) sets of commands were used to process, sort, and index those alignments made by bowtie2 (v2.3.4.3, 64bit; (Langmead and Salzberg, 2012); with parameters $k = 5$, $L = 12$, and "sensitive-local" switch), mapping the PE/SG reads against the SARS-CoV-2 reference genome (GenBank entry: NC_045512.2) and ensuring that the stored alignments were position sorted on the final bam files. Trinity (r20190503git; (Grabherr et al., 2013; Haas et al., 2013); min contig length = 100, k-mer size 31) produced contigs that later on were mapped over the reference genome obtaining the final scaffolds, then manually curated to finish the sequences. Due to the low number of reads recovered from wastewater sample, the corresponding assembly was filled with 93% of Ns to place the contigs into the assembled scaffold (this sequence is provided as Supplementary File 1). Mafft (v7.407; (Katoh and Standley, 2013); with localpair switch on) was chosen to calculate the multiple sequence alignment shown on Fig. 4, comparing the wastewater scaffold against a randomly sampled set of 600 Spanish and 600 international sequences from GISAID database (available on August 4th 2020; (Elbe and Buckland-Merrett, 2017; Shu and McCauley, 2017); see also Acknowledgments).

3. Results and discussion

3.1. SARS-CoV-2 quantification in clinical and wastewater samples

Viral SARS-CoV-2 copies were quantified in composite raw sewage samples collected in Barcelona different times of COVID-19 pandemic, March 2020, one week after the declaration of a 'state of alarm' in the country, and July 2020, 2020 just before the beginning of the second COVID-19 peak. The (RT)-qPCR were performed in quadruplicate with the direct NA extractions and in quadruplicate with 1/10 dilution of NA extractions, in order to avoid potential inhibitory effects. Concentration values obtained were $3.92\text{E}6$ (± 0.154) GC/L for N1 and $2.71\text{E}6$ (± 0.117) GC/L for N2 in March 2020 and $3.80\text{E}4$ (± 0.06) GC/L for N1 and $1.25\text{E}4$ (± 0.017) GC/L for N2 in July 2020. The concentrations for the human fecal indicator JCPyV were $1.99\text{E}5$ (± 0.057) GC/L and $8.73\text{E}5$ (± 0.026) GC/L in March and July respectively (Bofill-Mas et al., 2006). Enzymatic inhibition was not observed. These values are in accordance with the ones observed in the same WTP after a full year of weekly monitoring surveillance and 2 peak registered periods (Rusiñol et al., 2021).

Regarding the two SARS-CoV-2 laboratory-confirmed clinical samples from HUVH, the E gene cycle-threshold values were 15.5 and 16.4 from EPI_ISL_418860 and EPI_ISL_418861, respectively.

3.2. Urban wastewater virome during COVID-19 pandemic

Sewage samples from March and July were mass-sequenced in parallel using high throughput sequencing and probe-based target enrichment sequencing. The virome obtained in March, using target enrichment, resulted in almost 1 million reads belonging to 27 viral families. The distribution of the number of reads obtained for each vertebrate viral family is shown in Fig. 1. As expected, a wide variety of viral families infecting vertebrates was observed when applying the probed-

based capture methodology. This approach was successfully applied in the past to improve the deep sequencing in human-focused virome studies (Briese et al., 2015; Hjelmso et al., 2019; Martínez-Puchol et al., 2020), which is a key point in the study of viral species that are present in a low concentration in environmental and clinical samples. Traditional virome studies have shown that the most abundant viral sequences belong to bacteriophages and plant viruses (Cantalupo et al., 2011; Fernandez-Cassi et al., 2018), but depending on the sequencing platform used or the number of samples multiplexed, the total number of reads obtained could be reduced and relevant human and animal viral reads could go undetected. Target enrichment methods could overcome these limitations enabling the possibility of viral discovery.

In this study we used the VirCapSeq Enrichment Kit (Roche) intended to capture vertebrate viruses from complex samples prior to metagenomic sequencing, trying to improve the limitations of high throughput sequencing. It employs approximately 2 million biotinylated oligonucleotide probes designed to bind to the coding sequences of all viral taxa known to infect vertebrates at intervals of 50–100 nt. Libraries prepared from random primed cDNA are hybridized with the biotinylated probes and trapped with streptavidin magnetic beads. After magnetic capture and washing, NA are released from the beads and subjected to post-hybridization PCR prior to sequencing. In silico studies of the VirCapSeq panel performed in this study showed that although the platform was designed before the emergence of SARS-CoV-2, the kit nonetheless contained 21,414 fragments from 1838 sequences described for 346 *Coronaviridae* species. NCBI-BLASTn of the full sequences and the probes was run, with default parameters apart from the e-value set at $10\text{e}-25$, against SARS-CoV-2 reference genome (RefSeq ID: NC_045512; Wu et al., 2020a, 2020b). Among the 1838 full targets already contained in the VirCapSeq kit, 277 were found on 456 alignment hits, and 29 of those had hits above 90% identity; on the other hand, 104 probe fragments from 28 different sequences returned 104 hits.

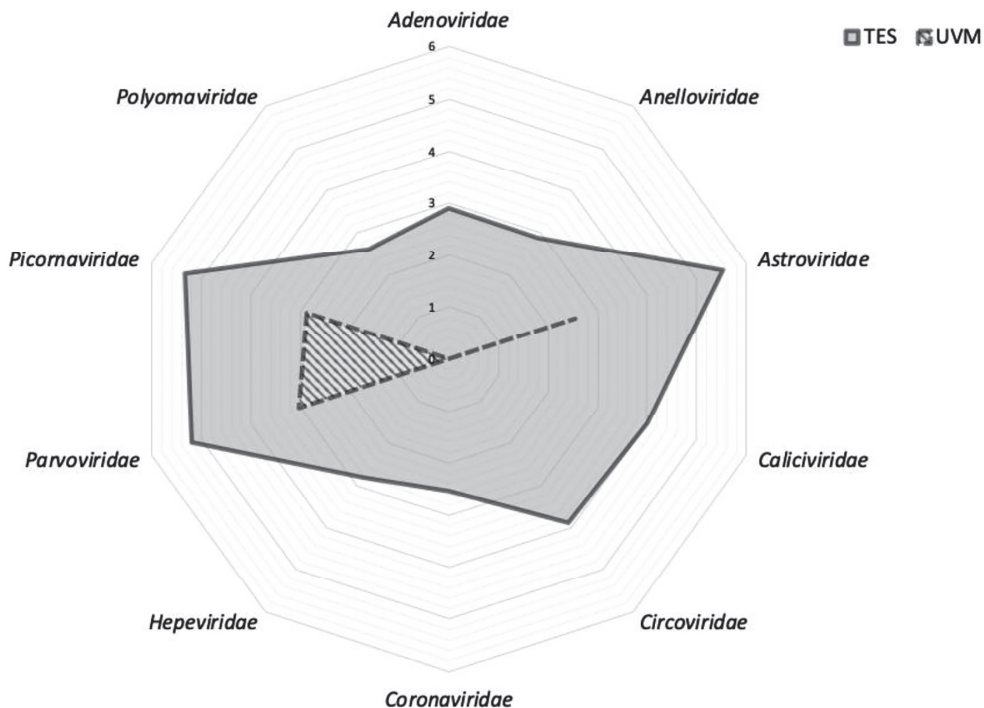


Fig. 1. Diversity and relative abundance of vertebrate viral families reads obtained from March sample. The results for untargeted viral metagenomics (UVM) are represented in dashed lines, regular lines represent the results after Target Enrichment NGS (TES). Relative abundance of reads is represented in log10.

Table 1
Coronaviridae family contigs description after performing Probe-capture targeted NGS in one urban wastewater sample (March 2020).

<i>Coronaviridae</i> species	Host	Reads	Contigs	Length (bp)	Region	Nt ID (%)	AA ID (%)
SARS-CoV-2	Human	45	8	203	ORF1ab polyprotein	100	100
				217	ORF1ab polyprotein	100	100
				126	ORF1ab polyprotein	100	100
				126	ORF1ab polyprotein	100	100
				178	ORF1ab polyprotein	100	100
				266	ORF3 protein	100	100
				516	Nucleocapsid protein	99.6	99.1
				399	Nucleocapsid protein	99.8	99.3
Betacoronavirus 1 (HCoV OC43)	Human	11	1	235	2'-O-methyltransferase	99.5	98.7
Feline coronavirus	Other vertebrates	33	1	524	ORF1ab polyprotein	91.8	86.5
				354	ORF1ab polyprotein	97.4	97.1
Lucheng Rn rat coronavirus	Other vertebrates	233	5	416	ORF1ab polyprotein	97.0	99.3
				542	ORF1ab polyprotein	97.0	99.4
				848	ORF1ab polyprotein	96.6	96.6
				1129	ORF1ab polyprotein	97.3	99.4
Canine coronavirus	Other vertebrates	14	1	231	ORF1ab polyprotein	96.1	94.8

The most abundant families obtained in March sample using target enrichment were *Astroviridae*, *Picornaviridae* and *Parvoviridae*. A complete list of all the identified viral species and families is presented in the Supplementary File 2. Homology searches against known viral genomes database retrieved *Coronaviridae* sequences listed in Table 1. One contig of a Betacoronavirus 1, typed as HCoV-OC43, was mapped with a high identity (99.5% at nucleotide level, 98.7% at aminoacid level) to the CoV 2'-O-methyltransferase, an enzyme that enables the mRNA cap formation, essential for viral RNA stability (Krafcikova et al., 2020). HCoV-OC43, one of the four seasonal HCoVs that are known to cause the common cold in humans, is especially prevalent during the winter months (Van der Hoek, 2015), and is known to have rodents as natural hosts and bovines as intermediate ones (Ye et al., 2020).

Eight SARS-CoV-2 contigs were obtained, with a total of 2.03 Kb representing 6.8% of the genome. Five contigs corresponded to the ORF1ab polyprotein region and one contig to the ORF3 protein, all of

them with 100% of nucleotide and amino acid identity with the genome Reference Sequence NC_045512. Additionally, two contigs, with sizes 516 bp and 316 bp, mapped against the nucleocapsid protein, separated with a gap between them of 231 bp, and presented a nucleotide identity of 99.6% and 99.8% to the reference sequence.

Regarding other members of the *Coronaviridae* family, one feline CoV contig (524 bp) matching the replicase ORF1ab polyprotein region was obtained in the analysis of sequences from the March wastewater sample. This virus causes asymptomatic persistent enteric infections in a high percentage of household and catteries' cats (Vogel et al., 2010). While feline CoV is highly prevalent in their hosts, its survival in sewage has been reported to be limited (Gundy et al., 2009). One contig from the same region (231 bp) was typed as canine CoV, known for being responsible of mild or moderate enteritis in dogs of all breeds and ages and with a high establishment in the environment (Pratelli, 2006). Finally, five contigs of Lucheng Rn rat CoV, comprising 3.28 Kb of the ORF1ab polyprotein, were obtained.

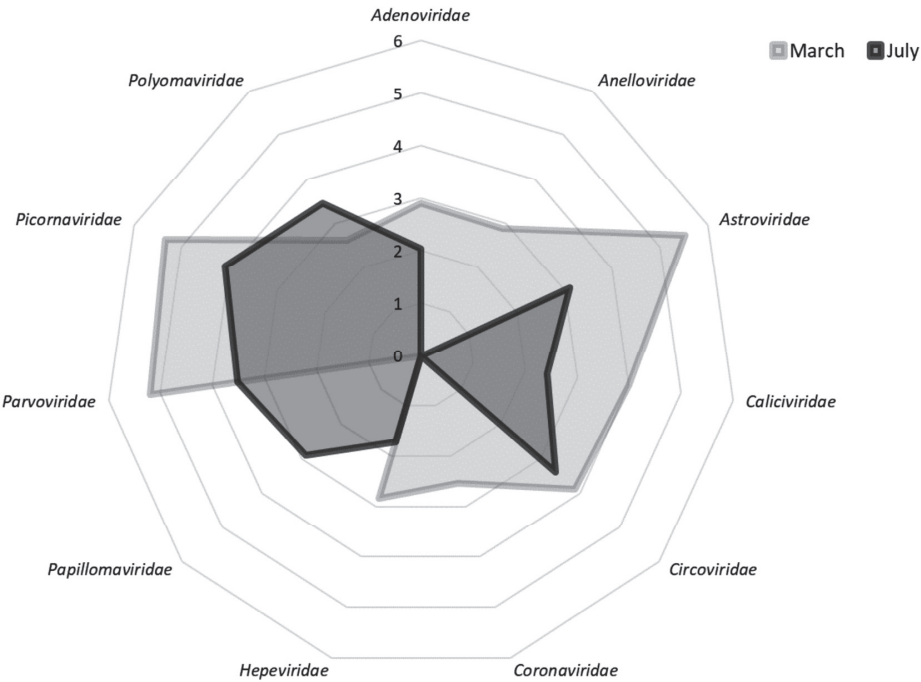


Fig. 2. Diversity and relative abundance of vertebrate viral families reads for March and July samples using Target Enrichment NGS. Relative abundance of reads is represented in log10.

This Alphacoronavirus was firstly described in 2014 in rat samples from China (Wang et al., 2015). At the time, rodents were not considered to be a relevant CoV reservoir, however it is now clear that this group of animals are in fact important reservoirs for alphacoronaviruses and also betacoronaviruses (Wartecki and Rzymiski, 2020).

New CoV members in bats has been a topic of interest within the scientific community due to the fact that these viruses could be the origin of important outbreaks. In this field, the use of metagenomics is important for a better understanding of evolution, epidemiology, and host-relationships of zoonotic and human viruses (Kivistö et al., 2020). These approaches made possible the discovery of new alphacoronaviruses (De Sabato et al., 2019; Hall et al., 2014), and recently the description of a betacoronavirus closely related to SARS-CoV-2 in *Rhinolophus* bats from China (Zhou et al., 2020). The low sensitivity of high throughput techniques or the lack of sequencing depth in samples with a high viral load has revealed enrichment NGS as a successful strategy for CoV surveillance and discovery in Asia (Li et al., 2020; Lim et al., 2019). Thus, NGS for CoV discovery should ideally be wide enough to detect unknown viruses but targeting viruses belonging to the family of interest among many other present in the analyzed samples.

Regarding the wastewater sample from July, its virome composition showed a lower proportion for almost all vertebrate viral families, except for *Polyomaviridae* and *Papillomaviridae*, compared with the sample from March (Fig. 2), members from the *Coronaviridae* family where no detected. The absence of SARS-CoV-2 could be due to the fact that COVID-19 incidence recorded when this sample was collected was low (175.2 cases/100,000 inhabitants), with wastewater quantifications of SARS-

CoV-2 two orders of magnitude lower than in March. The absence of HCoV-OC43 and animal CoV reads in the July sample could be explained by winter seasonality or by the fact that the transmission of these viruses could have been reduced due to the massive use of masks and the increase in hand washing from March to July as has been observed for other viruses.

3.3. Comparison of wastewater strains vs. clinical strains

SARS-CoV-2 contigs retrieved after applying targeted NGS to the March wastewater sample were compared with sequences obtained by applying the VirCapSeq enrichment approach to two clinical samples isolated in the same period. A short summary of the sequencing results for the three samples is provided in Supplementary File 2 (see Fig. 3 for an overview of the coverage distribution of the reads mapped over reference genome).

From a total of 5,006,516 paired-end reads, only 8 aligned uniquely at a single location and 20 aligned more than one time over the reference genome for the wastewater sample. On the contrary, VirCapSeq clinical samples starting with less clean paired-end reads, 375,769 and 316,398 for samples A and B, ended up with 1042 and 1856 uniquely mapped reads, respectively. The final scaffolds assembled from those sets of reads recovered up to 6.79% for wastewater sample, 88.38% for clinical sample A, and 89.83% for clinical sample B, of the 29,903 bp of reference SARS-CoV-2 genome sequence.

Other studies have reported the use of VirCapSeq and Twist Bioscience panels for sequencing of clinical samples, determining the

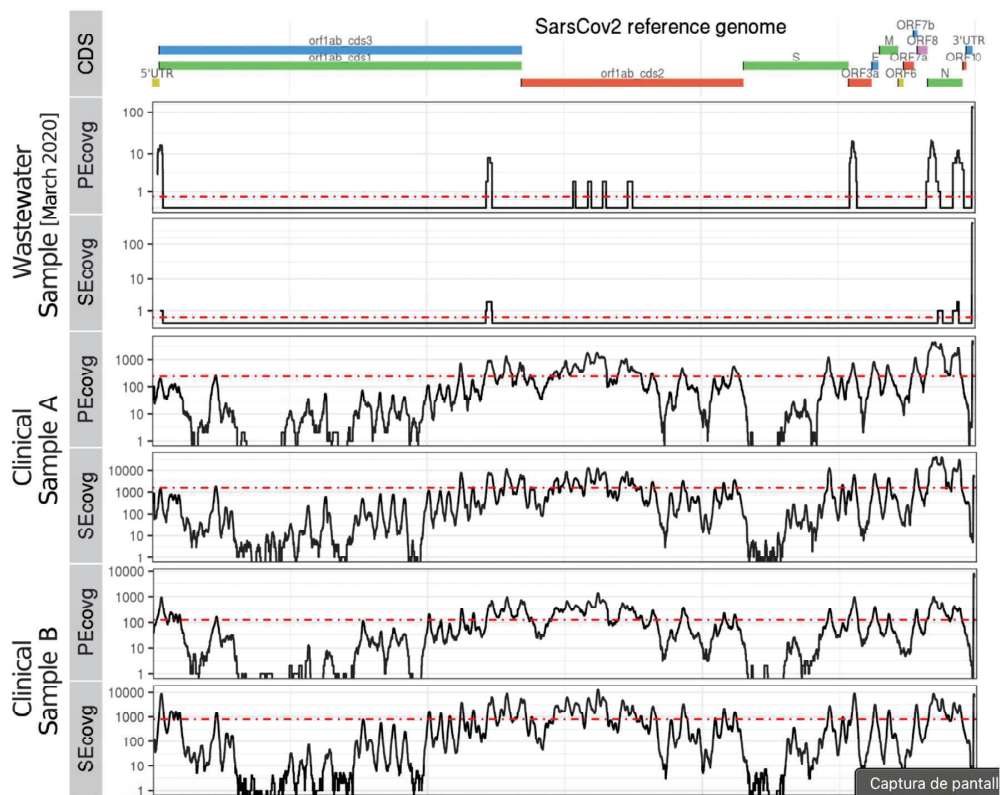


Fig. 3. Summary of reads coverage along the SARS-CoV-2 reference genome sequence. The number of reads per position for the clean Pair-End reads (PEcovg) and for the Single-End reads (SEcovg) is shown here for the three samples. SE-reads were produced when cleaning the raw PE-reads, corresponding to the single member of a pair or reads that passed the cleaning and trimming criteria while the other read was discarded. Top track represents the segments defining the open reading frames encoding for the viral proteins (except the trailing 5'UTR and 3'UTR segments that are not protein-coding and were only included for illustrative purposes). The dashed red line represents the average coverage. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

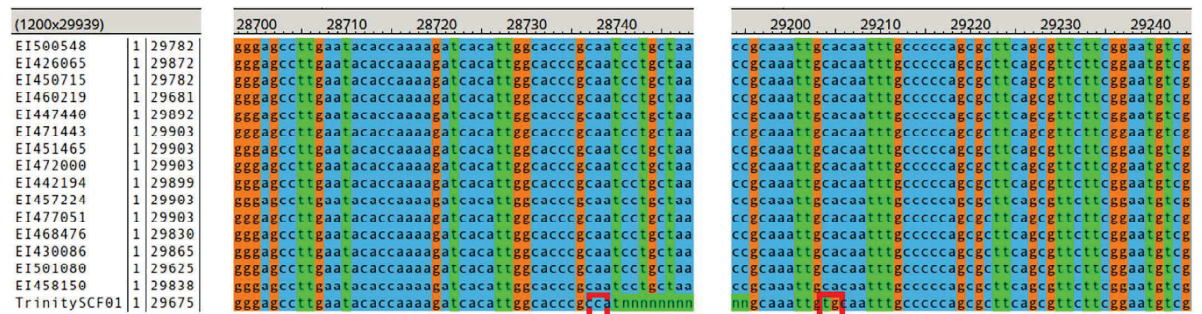


Fig. 4. Nucleotide mismatches found in the nucleocapsid protein region. The red boxes highlight the 3 nucleotides that differ on the sequence reconstructed from the recovered reads with respect to the alignment with 1200 randomly chosen GISAID SARS-CoV-2 sequences (including the reference). Those nucleotide substitutions correspond to 1 synonymous and 2 non-synonymous changes: 28721.caa[Q] <> 28,514.caa[P], 29,187.tgc[C] == 28,959.tgt[C], 29,188.aca[T] <> 28,960.gca[A]. Those three positions do not change on the alignment for all the other provided sequences yet are supported by only three reads which in fact does not allow defining them as variants. Scale on top corresponds to the positions relative to the full alignment (29,939 columns) that is available as Supplementary File 4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

importance of having low qPCR Ct values in a given sample for obtaining a high genome coverage after conducting NGS (Carbo et al., 2020; Klempt et al., 2020).

SARS-CoV-2 sequences obtained from sewage were compared with other 1200 sequences obtained from clinical isolates available on the GISAID database (Elbe and Buckland-Merrett, 2017) (<https://www.gisaid.org/epiflu-applications/hcov-19-genomic-epidemiology/>) to elucidate the presence of viral variants. While a high degree of similarity was found for the assembled contigs, 3 nucleotide mismatches were identified within the nucleocapsid protein region (Fig. 4). These divergences could represent single nucleotide variants (SNV), but they were only supported by 3 overlapping reads. More read coverage would be needed in this region to further assess these differences, although it is worth noting that those changes were not described in clinical samples at the moment. Amplicon sequencing panels covering the totality of the genome are available now for studying SARS-CoV-2 genetic variants. These approaches, as well as specific SARS-CoV-2 and Human CoV capture panels, would have enabled the acquisition of sufficient genome coverage and consequently a better SNV typing (Nasir et al., 2020; Xiao et al., 2020) than other targeted approaches directed towards wider groups of viruses, like VirCapSeq capture used in this study. Even so, NGS methods still could not compete with other molecular approaches, as RT-qPCR, as both the time to obtain results and the processing price are higher. In contrast, these sequencing-based methodologies may be more sensitive than RT-qPCR (Charlebois et al., 2020) and give broader information about genetic variants and characteristics of other members from the same family that could not be detected by a single primer amplification-based method. This could be the case of Amplicon Deep Sequencing approaches based on massive sequencing of a PCR produced amplicon. High throughput NGS approaches have as a main advantage that allow detection of unknown viral sequences, in contrast, PCR-based detection relies of previous knowledge and primer design of viral targets. RT-qPCR and NGS-based methods do not necessarily compete but complement each other to provide useful information in terms of wastewater-based epidemiology. Moreover, NGS methods are in constant development, thus it is expected that in a short period of time could compete in terms of time and cost with other molecular approaches.

4. Conclusions

- The application of a targeted NGS has provided nearly the whole genome of SARS-CoV-2 from two clinical samples and eight SARS-CoV-2 contigs from one wastewater sample, both type of samples obtained in March 2020 from the same geographical area.

- Because of low genome coverage obtained in the wastewater sample, the 3 nucleotide differences observed among environmental and clinical samples could not be further assessed but results obtained show a high degree of similarity between environmental and clinical samples.
- The results of the excreted virome in wastewater showed that there is co-circulation, in urban areas, of human and animal coronaviruses infecting domestic animals and rodents.
- The use of a Target Enrichment panel designed to cover vertebrate viruses allowed the acquisition of SARS-CoV-2 sequences as part of the wastewater virome within the context of a COVID-19 pandemic.
- Specific SARS-CoV-2/human CoV panels should be used for studying SARS-CoV-2 and other Human CoV genetic diversity while panels targeting a wide variety of viral families would be a better choice for evaluating the co-circulation of known and unknown human and animal CoV, which may be of relevance regarding potentially zoonosis and viral discovery.
- SARS-CoV-2 can be now considered a new member of the sewage virome. Its presence in this type of samples after global vaccination campaigns should be further evaluated as well as the presence of other human CoV that could change their circulation patterns due to immune cross reactivity phenomena.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scitotenv.2021.149562>.

Availability of data and material

The NGS datasets generated during the current study are available in zenodo under the DOI number <https://doi.org/10.5281/zenodo.4252765>; raw sequencing reads for all the analyzed samples can be also accessed from NCBI-SRA Bioproject PRJNA689221.

Funding

This study was partially supported by the Ministry of Science, Innovation and Universities (AGL2017-86797-C2-1-R) through the University of Barcelona and the Direcció General de Recerca i Innovació en Salut (DGRIS) Catalan Health Ministry Generalitat de Catalunya through Vall d'Hebron Research Institute (VHIR). Sílvia Bofill-Mas is a Serra-Hunter fellow at the University of Barcelona.

CRediT authorship contribution statement

Conceptualization, S.B.-M. and S.M.-P.; methodology, S.M.-P., M.I., M. R., E.F., C.M.; software, J.A., S.M.-P.; validation, S.B.-M., R.G. and J.A.; formal

analysis, S.M—P, S.B-M and J.A; investigation, S.M-P and S.B-M.; resources, S.B-M, R.G, J.Q, A.A.; data curation, S.M-P and C.A; writing—original draft preparation, S.M-P and S.B-M.; writing—review and editing, S. M—P, S.B-M., M.I, M.R, E.F, C.M, C.A, A.A, J.Q, R.G a J.A; visualization, S.M-P and J.A.; supervision, S.B-M.; project administration, S.B-M, R.G, A.A, J.Q.; funding acquisition, S.B-M, R.G, A.A, J.Q. All authors have read and agreed to the published version of the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the members of ICRA and the sewage treatment plants personnel for the coordination of the wastewater sampling. We also gratefully acknowledge the authors originating and submitting laboratories of the SARS-CoV-2 sequences from GISAID's EpiCov Database (<https://www.epicov.org/epi3/>) used to build the Fig. 4 alignment; access to the individual isolates is facilitated through GISAID web site (<https://www.gisaid.org/>).

References

- Barnett, D.W., Garrison, E.K., Quinlan, A.R., Strömberg, M.P., Marth, G.T., 2011. Bamtools: a C API and toolkit for analyzing and managing BAM files. *Bioinformatics* 27, 1691–1692. <https://doi.org/10.1093/bioinformatics/btr174>.
- Bofill-Mas, S., Albinana-Gimenez, N., Clemente-Casares, P., Hundesa, A., Rodriguez-Manzano, J., Allard, A., Calvo, M., Girones, R., 2006. Quantification and stability of human adenoviruses and polyomavirus JCPyV in wastewater matrices. *Appl. Environ. Microbiol.* <https://doi.org/10.1128/AEM.00965-06>.
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Briese, T., Kapoor, A., Mishra, N., Jain, K., Kumar, A., Jabado, O.J., Ian Lipkina, W., 2015. Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *MBio* 6, 1–12. <https://doi.org/10.1128/mBio.01491-15>.
- Cantalupo, P.G., Calgua, B., Zhao, G., Hundesa, A., Wier, A.D., Katz, J.P., Grabe, M., Hendrix, R.W., Girones, R., Wang, D., Pipas, J.M., 2011. Raw sewage harbors diverse viral populations. *MBio* 2, <https://doi.org/10.1128/mBio.00180-11> e00180-11.
- Carbo, E.C., Sidorov, I.A., Zevenhoven-Dobbe, J.C., Snijder, E.J., Claas, E.C., Laros, J.F.J., Kroes, A.C.M., de Vries, J.J.C., 2020. Coronavirus discovery by metagenomic sequencing: a tool for pandemic preparedness. *J. Clin. Virol.* 131, 104594. <https://doi.org/10.1016/j.jcv.2020.104594>.
- Charlebois, R.L., Sathiamoorthy, S., Logvinoff, C., Gissoni-Lex, L., Mallet, L., Ng, S.H.S., 2020. Sensitivity and breadth of detection of high-throughput sequencing for adventitious virus detection. *npj Vaccines* 5, 1–8. <https://doi.org/10.1038/s41541-020-0207-4>.
- De Sabato, L., Lelli, D., Faccin, F., Canziani, S., Di Bartolo, I., Vaccari, G., Moreno, A., 2019. Full genome characterization of two novel alpha-coronavirus species from Italian bats. *Virus Res.* 260, 60–66. <https://doi.org/10.1016/j.virusres.2018.11.007>.
- Elbe, S., Buckland-Merrett, G., 2017. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob. Chall.* 1, 33–46. <https://doi.org/10.1002/gch2.1018>.
- Fernandez-Cassi, X., Timoneda, N., Martínez-Puchol, S., Rusiñol, M., Rodríguez-Manzano, J., Figuerola, N., Bofill-Mas, S., Abril, J.F., Girones, R., 2018. Metagenomics for the study of viruses in urban sewage as a tool for public health surveillance. *Sci. Total Environ.* 618, 870–880. <https://doi.org/10.1016/j.scitotenv.2017.08.249>.
- Grabherr, M.G., Haas, Brian J., Joshua, Moran Yassour, Levin, Z., Thompson, Dawn A., Amit, Ido, Adiconis, Xian, Fan, Lin, Raychowdhury, Raktima, Zeng, Qiangdong, Chen, Zehua, Mauceli, Evan, Hacohen, Nir, Gnirke, Andreas, Rhind, Nicholas, di Palma, Federica, Bruce, W., N. and A.R., Friedman, 2013. Trinity: reconstructing a full-length transcriptome without a genome from RNA-seq data. *Nat. Biotechnol.* 29, 644–652. <https://doi.org/10.1038/nbt.1883>.
- Gundy, P.M., Gerba, C.P., Pepper, I.L., 2009. Survival of coronaviruses in water and wastewater. *Food Environ. Virol.* 1, 10–14. <https://doi.org/10.1007/s12560-008-9001-6>.
- Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Phillip, D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Macmanes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., Williams, T., Dewey, C.N., Henschel, R., Leduc, R.D., Friedman, N., Regev, A., 2013. De novo transcript sequence reconstruction from RNA-seq: reference generation and analysis with trinity. *Nat. Protoc.* <https://doi.org/10.1038/nprot.2013.084>.
- Hall, R.J., Wang, J., Peacey, M., Moore, N.E., McInnes, K., Tompkins, D.M., 2014. New alphacoronavirus in *Mystacina tuberculata* bats, New Zealand. *Emerg. Infect. Dis.* 20, 697–700. <https://doi.org/10.3201/eid2004.131441>.
- Hjelm, M.H., Møllerup, S., Jensen, R.H., Pietroni, C., Lukjancenko, O., Schultz, A.C., Aarestrup, F.M., Hansen, A.J., 2019. Metagenomic analysis of viruses in toilet waste from long distance flights—A new procedure for global infectious disease surveillance. *PLoS One* 14, e0210368. <https://doi.org/10.1371/journal.pone.0210368>.
- Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. <https://doi.org/10.1093/molbev/mst010>.
- Kivistö, I., Tidenberg, E.M., Lilley, T., Suominen, K., Forbes, K.M., Vapalahti, O., Huovilainen, A., Sironen, T., 2020. First report of coronaviruses in northern European bats. *Vector-Borne Zoonotic Dis.* 20, 155–158. <https://doi.org/10.1089/vbz.2018.2367>.
- Klemp, P., Brož, P., Kašný, M., Novotný, A., Kvapilová, K., Kvapil, P., 2020. Performance of targeted library preparation solutions for SARS-CoV-2 whole genome analysis. *Diagnostics* 10, 769. <https://doi.org/10.3390/diagnostics10100769>.
- Krafcikova, P., Silhan, J., Nencka, R., Boura, E., 2020. Structural analysis of the SARS-CoV-2 methyltransferase complex involved in RNA cap creation bound to sinefungin. *Nat. Commun.* 11, 1–7. <https://doi.org/10.1038/s41467-020-17495-9>.
- Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with bowtie 2. *Nat. Methods* 9, 357–359. <https://doi.org/10.1038/nmeth.1923>.
- Li, B., Si, H.-R., Zhu, Y., Yang, X.-L., Anderson, D.E., Shi, Z.-L., Wang, L.-F., Zhou, P., 2020. Discovery of bat coronaviruses through surveillance and probe capture-based next-generation sequencing. *mSphere* 5, 1–10. <https://doi.org/10.1128/mSphere.00807-19>.
- Lim, X.F., Lee, C.B., Pascoe, S.M., How, C.B., Chan, S., Tan, J.H., Yang, X., Zhou, P., Shi, Z., Sessions, O.M., Wang, L.F., Ng, L.C., Anderson, D.E., Yap, G., 2019. Detection and characterization of a novel bat-borne coronavirus in Singapore using multiple molecular approaches. *J. Gen. Virol.* 100, 1363–1374. <https://doi.org/10.1099/jgv.0.001307>.
- Martínez-Puchol, S., Rusiñol, M., Fernández-Cassi, X., Timoneda, N., Itarte, M., Andrés, C., Antón, A., Abril, J.F., Girones, R., Bofill-Mas, S., 2020. Characterisation of the sewage virome: comparison of NGS tools and occurrence of significant pathogens. *Sci. Total Environ.* 713. <https://doi.org/10.1016/j.scitotenv.2020.136604>.
- Nasir, J.A., Kozak, R.A., Aftanas, P., Raphenya, A.R., Smith, K.M., Maguire, F., Maan, H., Alruwaili, M., Banerjee, A., Mbareche, H., Alcock, B.P., Knox, N.C., Mossman, K., Wang, B., Hiscio, J.A., McArthur, A.G., Mubareka, S., 2020. A comparison of whole genome sequencing of SARS-CoV-2 using amplicon-based sequencing, random hexamers, and bait capture. *Viruses* 12. <https://doi.org/10.3390/v12080895>.
- Ng, T.F.F., Marine, R., Wang, C., Simmonds, P., Kapusinsky, B., Bodhidatta, L., Oderinde, B.S., Wommack, K.E., Delwart, E., 2012. High variety of known and new RNA and DNA viruses of diverse origins in untreated sewage. *J. Virol.* 86, 12161–12175. <https://doi.org/10.1128/JVI.00869-12>.
- Pratelli, A., 2006. Genetic evolution of canine coronavirus and recent advances in prophylaxis. *Vet. Res.* 37, 191–200. <https://doi.org/10.1051/vetres:2005053>.
- Rusiñol, M., Zammit, I., Itarte, M., Forés, E., Martínez-Puchol, S., Girones, R., Borrego, C., Corominas, L., Bofill-Mas, S., 2021. Monitoring waves of the COVID-19 pandemic: inferences from WWTPs of different sizes. *Sci. Total Environ.* 787, 147463. <https://doi.org/10.1016/j.scitotenv.2021.147463>.
- Shu, Y., McCauley, J., 2017. GISAID: global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance* 22, 2–4. <https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494>.
- Van der Hoek, L., 2015. Human coronaviruses: what do they cause? *Antivir. Ther.* 12.
- Vilsker, M., Moosa, Y., Nooij, S., Fonseca, V., Ghysens, Y., Dumon, C., Pauwels, R., Alcantara, L.C., Vanden Eynden, E., Vandamme, A.-M., Deforche, K., de Oliveira, T., 2018. Genome detective: an automated system for virus identification from high-throughput sequencing data. *Bioinformatics* 1–3. <https://doi.org/10.1093/bioinformatics/bty695>.
- Vogel, L., Van Der Lubben, M., Te Lintelo, E.G., Bekker, C.P.J., Geerts, T., Schuijff, L.S., Grinwis, G.C.M., Egberink, H.F., Rottier, P.J.M., 2010. Pathogenic characteristics of persistent feline enteric coronavirus infection in cats. *Vet. Res.* 41. <https://doi.org/10.1051/vetres/2010043>.
- Wang, W., Lin, X., Guo, W., Zhou, R., Wang, M., Wang, C.-Q., Ge, S., Mei, S.-H., Li, M.-H., Shi, M., Holmes, E.C., Zhang, Y.-Z., 2015. Discovery, diversity and evolution of novel coronaviruses sampled from rodents in China. *Virology* 474, 19–27. <https://doi.org/10.1016/j.virol.2014.10.017>.
- Wartecki, A., Rzymiski, P., 2020. On the coronaviruses and their associations with the aquatic environment and wastewater. *Water (Switzerland)* 12, 1–27. <https://doi.org/10.3390/W12061598>.
- Wölfel, R., Corman, V.M., Guggemos, W., Seilmaier, M., Zange, S., Müller, M.A., Niemeyer, D., Jones, T.C., Vollmar, P., Rothe, C., Hoelscher, M., Bleicker, T., Brünink, S., Schneider, J., Ehmman, R., Zwirgmaier, K., Drosten, C., Wendtner, C., 2020. Virological assessment of hospitalized patients with COVID-2019. *Nature* 581, 465–469. <https://doi.org/10.1038/s41586-020-2196-x>.
- Wu, F., Zhao, S., Yu, B., Chen, Y.M., Wang, W., Song, Z.G., Hu, Y., Tao, Z.W., Tian, J.H., Pei, Y.Y., Yuan, M.L., Zhang, Y.L., Dai, F.H., Liu, Y., Wang, Q.M., Zheng, J.J., Xu, L., Holmes, E.C., Zhang, Y.Z., 2020. A new coronavirus associated with human respiratory disease in China. *Nature* 579, 265–269. <https://doi.org/10.1038/s41586-020-008-3>.
- Wu, Y., Guo, C., Tang, L., Hong, Z., Zhou, J., Dong, X., Yin, H., Xiao, Q., Tang, Y., Qu, X., Kuang, L., Fang, X., Mishra, N., Lu, J., Shan, H., Jiang, G., Huang, X., 2020. Prolonged presence of SARS-CoV-2 viral RNA in faecal samples. *Lancet Gastroenterol. Hepatol.* 5, 434–435. [https://doi.org/10.1016/S2468-1253\(20\)30083-2](https://doi.org/10.1016/S2468-1253(20)30083-2).
- Xiao, M., Liu, X., Ji, J., Li, M., Li, J., Li, J., Li, J., Sun, W., Ren, P., Yang, G., Zhao, J., Liang, T., Ren, H., Chen, T., Zhong, H., Song, W., Wang, Y., Deng, Z., Zhao, Y., Ou, Z., Wang, D., Cai, J., Cheng, X., Feng, T., Wu, H., Gong, Y., Yang, H., Wang, J., Xu, X., Zhu, S., Chen, F., Zhang, Y., Chen, W., Li, Y., Li, Junhua, 2020. Multiple approaches for massively parallel sequencing of SARS-CoV-2 genomes directly from clinical samples. *Genome Med.* 12, 1–15. <https://doi.org/10.1186/s13073-020-00751-4>.
- Ye, Z.W., Yuan, S., Yuen, K.S., Fung, S.Y., Chan, C.P., Jin, D.Y., 2020. Zoonotic origins of human coronaviruses. *Int. J. Biol. Sci.* 16, 1686–1697. <https://doi.org/10.7150/ijbs.45472>.
- Zhou, H., Chen, X., Hu, T., Li, J., Song, H., Liu, Y., Wang, P., Liu, D., Yang, J., Holmes, E.C., Hughes, A.C., Bi, Y., Shi, W., 2020. A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein. *Curr. Biol.* 30, 2196–2203.e3. <https://doi.org/10.1016/j.cub.2020.05.023>.

Article 3. Metaviròmica amb enriquiment de dianes per a la vigilància de coronavirus en aigües residuals i mostres d'origen animal

“Target enrichment metaviromics enables comprehensive surveillance of coronaviruses in environmental and animal samples”

Martínez-Puchol, S., Tarradas-Alemany, M., Mejías-Molina, C., Itarte, M., Rusiñol, M., Baliellas, J., Abasolo, N., Canela, N., Monastiri, A., López-Roig, M., Serra-Cobo, J., Abril, J. F. i Bofill-Mas, S.

Article en procés de revisió a la revista *Heliyon*.

La pandèmia de la COVID-19 ha posat de manifest la importància d'entendre el paper dels animals en la transmissió dels coronavirus i el seu impacte en la salut humana. Per a un control efectiu d'aquests virus, és essencial tenir un enfocament *One Health* que integri la salut humana, animal i ambiental. Les tècniques de seqüenciació de NGS han demostrat ser útils en la identificació i el monitoratge de l'evolució de coronavirus emergents, com ara el SARS-CoV-2. No obstant això, estudiar l'ocurrència i la diversitat vírica en mostres ambientals i animals suposa un repte, degut a la complexitat de les comunitats víriques i la baixa concentració de virus en aquest tipus de mostres.

Com s'ha vist en l'article anterior, la seqüenciació amb enriquiment de dianes ha demostrat ser una estratègia eficaç per investigar virus presents en mostres complexes. Aquesta tècnica permet millorar l'eficiència de la detecció i la caracterització de virus mitjançant la captura i l'enriquiment de genomes vírics amb sondes específiques de seqüències de virus de vertebrats disponibles comercialment.

En aquest estudi, aquest enriquiment es va dur un pas més enllà, i es va desenvolupar i validar un panell de captura per a la seqüenciació massiva amb enriquiment de dianes dirigit a estudiar la diversitat de coronavirus presents en diferents mostres complexes d'origen ambiental i animal. Els resultats van demostrar l'eficàcia del panell per capturar i seqüenciar coronavirus en mostres d'aigua residual urbana, mostres de guano de ratpenat i lixiviats de cadàvers d'animals de granja. Amb aquest panell, es va millorar la seqüenciació de coronavirus en mostres d'aigua residual respecte el panell utilitzat en l'article anterior, VirCapSeq-VERT, obtenint així una gran diversitat de coronavirus, sobretot coronavirus que infecten humans i animals domèstics. En les mostres de guano de ratpenat es van identificar quatre espècies diferents de coronavirus que infecten ratpenat, i també es van detectar *reads* associats al coronavirus de civeta HKU8, el qual s'ha especulat que podria estar associat a un possible esdeveniment de transmissió entre espècies a partir d'un coronavirus de ratpenat. D'altra banda, de les mostres de lixiviats d'animals es va obtenir una diversitat vírica menor en comparació a la resta de mostres, identificant coronavirus que infecten aus, porcs, conills, bestiar boví i gossos, depenent el tipus de lixiviat analitzat.

La informació obtinguda a partir d'aquesta investigació pot contribuir significativament a la detecció primerenca, la vigilància i les mesures de control dels coronavirus, incloent-hi el descobriment de virus nous i la identificació de zoonosis potencials. A més, aquest panell de captura mostra la possibilitat d'estudiar altres famílies víriques d'interès i monitoritzar la diversitat vírica en diferents poblacions animals.

TARGET ENRICHMENT METAVIROMICS ENABLES COMPREHENSIVE SURVEILLANCE OF CORONAVIRUSES IN ENVIRONMENTAL AND ANIMAL SAMPLES

Sandra Martínez-Puchol^{1,7}, Maria Tarradas-Alemany^{2,5}, Cristina Mejías-Molina^{1,4}, Marta Itarte^{1,4},
Marta Rusiñol^{1,4}, Jordi Baliellas⁸, Nerea Abasolo⁹, Núria Canela⁹, Abir Monastiri^{3,6}, Marc López-
Roig^{3,6}, Jordi Serra-Cobo^{3,6}, Josep F. Abril^{2,5}, Sílvia Bofill-Mas^{1,4*}

¹ Laboratory of Viruses Contaminants of Water and Food; Departament de Genètica. Microbiologia i Estadística, Universitat de Barcelona. Barcelona, Catalonia, Spain.

² Computational Genomics Lab; Departament de Genètica. Microbiologia i Estadística, Universitat de Barcelona. Barcelona, Catalonia, Spain.

³ Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals, Universitat de Barcelona, Catalonia, Spain.

⁴ Institut de Recerca de l'Aigua (IdRA), Universitat de Barcelona, Barcelona, Catalonia, Spain.

⁵ Institut de Biomedicina (IBUB), Universitat de Barcelona, Barcelona, Catalonia, Spain.

⁶ Institut de Recerca de la Biodiversitat (IRBiO), Universitat de Barcelona, Catalonia, Spain.

⁷ Vicerectorat de Recerca. Universitat de Barcelona, Barcelona, Catalonia, Spain.

⁸ Grup Sanejament Porcí, Lleida, Catalonia, Spain.

⁹ Centre for Omic Sciences (COS), Joint Unit Universitat Rovira I Virgili-EURECAT, Unique Scientific and Technical Infrastructures (ICTS), Eurecat, Centre Tecnològic de Catalunya, Catalonia, Spain.

* Correspondence: sbofill@ub.edu

Abstract

The COVID-19 pandemic has underscored the importance of understanding the role of animals in the transmission of coronaviruses (CoVs) and their impact on human health. A One Health approach, integrating human, animal, and environmental health, is essential for effective CoVs control. Next-generation sequencing has played a pivotal role in identifying and monitoring the evolution of novel CoVs strains, like SARS-CoV-2. However, viral occurrence and diversity studies in environmental and animal samples are challenging because of the complexity of viral communities and low abundance of viruses in these samples.

Target enrichment sequencing (TES) has emerged as valuable tool for investigating viral families in challenging samples. This approach involves the specific capture and enrichment of viral genomes using sequence-specific probes, thereby enhancing the efficiency of detection and characterization. In this study, we aimed to develop and validate a TES panel to study CoVs in various complex environmental and animal derived samples. The results demonstrated the panel's effectiveness in capturing and sequencing a wide diversity of CoVs providing valuable insights into their abundance and host diversity in urban wastewater, farm animal corpses lixivates and bat guano samples. In sewage samples, CoVs were detected solely when TES was employed while in guano samples, sequencing of CoVs species was achieved in 2 out of 4 samples showing an almost three-logarithmic increase in the number of reads obtained in comparison with the untargeted approach. For animal lixivates, only the TES application enabled the acquisition of CoVs reads. The information obtained can significantly contribute to early detection, surveillance, and control measures for CoVs, including viral discovery and potential spillover events. Additionally, this sequencing panel shows potential for studying other significant viral families and monitoring viral diversity in different animal populations.

1. Introduction

The COVID-19 pandemic highlighted the necessity of understanding the role of animals in the transmission of coronaviruses (CoVs) and their potential impact on human health (Pekar et al., 2022). CoVs can cause a range of illnesses in infected humans and animals. Their transmission occurs through respiratory droplets, but other bodily fluids could be involved and for this, studying their presence in human and animal excreta can provide valuable information about CoVs circulation (Maryam et al., 2023). CoVs are zoonotic pathogens capable of cross-species transmission, resulting in outbreaks in both humans and animals. Consequently, the implementation of effective strategies for controlling CoVs requires adopting a One Health approach that integrates human, animal, and environmental health (Ruiz-Aravena et al., 2022; Sharun et al., 2021). Therefore, it becomes imperative to investigate the interactions between humans, animals, and their shared environments to identify potential reservoirs and intermediate CoVs hosts. This knowledge plays a crucial role in proactively preventing future spillover events and emerging infectious diseases, helping early detection of potential outbreaks and prompt response measures (Islam et al., 2022).

To identify potential sources of human infections or the emergence of new zoonotic species or strains, it is essential to study the genetic diversity and evolutionary dynamics of CoVs in both domesticated and wild animals, as well as in the environment (Carlson et al., 2022; Jo et al., 2021). Next-generation sequencing (NGS) has played a crucial role in identifying novel CoVs strains, including those responsible for recent outbreaks like SARS-CoV-2, and has enabled the monitoring of their evolution (Charre et al., 2020; Liu et al., 2022). This technology facilitates the detection of genetic mutations, and the tracking of the emergence and spread of new variants within populations.

Virome studies on environmental and animal samples present significant challenges due to the complex and diverse nature of viral communities within these samples. One of the primary obstacles in these studies lies in the extensive viral diversity observed in environmental samples (Cantalupo et al., 2011), which often remains unexplored or insufficiently characterized. Traditional virological methods often have limitations in capturing and identifying the full extent of viral genomes within a sample. Moreover, viruses may exhibit low abundance compared to other microorganisms, further complicating their detection and analysis (Fernandez-Cassi et al., 2018a).

To overcome these challenges, target enrichment sequencing (TES) has emerged as valuable tool, proving highly beneficial in the investigation of various viral families, including CoVs, in clinical and environmental samples. However, previous attempts using a commercial probe-panel directed to vertebrate viruses only yielded a few reads from SARS-CoV-2 and other human and animal CoVs in sewage (Martínez-Puchol et al., 2020; 2021). This technique enables the specific capture and enrichment of viral genomes using sequence-specific probes, thereby enhancing the sensitivity and efficiency of detection and characterization (Briese et al., 2015; Paskey et al., 2019). The use of targeted panels designed to capture specific viral families of interest within challenging samples can enhance NGS approaches, resulting in increased sequencing depth and providing a more comprehensive understanding of viral diversity. This becomes particularly crucial when studying reservoir animals such as bats (Zhou et al., 2020) or potential intermediate hosts like farm animals (Khamassi Khbou et al., 2021).

Consequently, the aim of this study was to develop and validate the efficacy of a TES panel specifically designed to capture members of the *Coronaviridae* family to study the diversity of CoVs in sewage, bat guano and farm-derived animal samples.

90

91 **2. Materials and Methods**

92 *Sample collection, viral concentration, and nucleic acid extraction*

93 Urban 24-hours composite sewage samples were collected on November 16th (Sewage02), and
94 December 14th (Sewage03), 2021 and on January 3rd, 2022 (Sewage01) from a wastewater treatment
95 plant (WWTP) located in the city of Barcelona. This plant serves up to 2.8 million population
96 equivalents and receives both domestic and industrial waste from the sewer system. Upon collection,
97 the samples were placed in a sterile container and kept at 4°C until viral particles were concentrated
98 from 100ml of sewage. In brief, the samples underwent initial debris removal through centrifugation
99 at 4750×g for 30 min and resulting supernatant (80mL) was ultrafiltered using the automatic
100 Concentration Pipette (CP-Select™) with 150kDa tips (Innovaprep, Missouri, USA) (Forés et al.,
101 2021). After the process, viral concentrates were recovered into a volume of 300µL. Nucleic acids
102 were extracted using using QIAamp RNAViral Mini Kit (Qiagen, Hilden, Germany).

103 Pools of guano were collected from bat roosts in four different locations in Catalonia (33) and Balearic
104 Islands (1). Refuge 1 (June 2019, guano 01) shelters various bat species, notably including significant
105 colonies of *Miniopterus schreibersii* and *Myotis myotis* of hundred individuals. Refuge 2 (March
106 2016, Guano 02) hosts a hibernating colony of over 2,000 *Miniopterus schreibersii*. Refuge 3 (August
107 2017, Guano 03) hosts various bat species, including significant breeding colonies of *Miniopterus*
108 *schreibersii* and *Myotis myotis* of hundred individuals. Refuge 4 (December 2015, Guano 04) hosts
109 a very important colony of 15,000-17,000 *Miniopterus schreibersii*. All guano samples were obtained
110 either before or after the hibernating or breeding periods and were preserved using RNAlater
111 (Invitrogen, Massachusetts, USA) and stored at -80°C.

112 From each guano pool, viral particles were concentrated from 1gr of feces using glycine (0.25N, 9.5
113 pH) and ultracentrifugation protocol described by Pina *et al.* (Pina et al., 1998). To eliminate free-
114 DNA from sewage and guano viral concentrates, Turbo DNase (Invitrogen, Massachusetts, USA)
115 was used, followed by nucleic acid extraction using QIAamp Viral RNA mini kit (Qiagen, Hilden,
116 Germany) with an initial concentrate volume of 280µL and an elution volume of 80µL.

117 Animal lixiviate samples from chicken, porcine and rabbit from animal corpse trucks were collected
118 using sterile containers directly from the leachate generated during discharge at a rendering plant

located in Catalonia and stored at 4 °C during transport to the laboratory where they were analyzed. Viruses were released from 2 gr of lixiviate by vortexing and then gentle shaking with 20ml of glycine buffer (0.25N, 9.5pH) for 20 min at room temperature. The mixture was then centrifuged at 8000xg for 15min. Total nucleic acids were purified from the supernatant using the Maxwell® Enviro Wastewater TNA Kit (Promega, Wisconsin, USA), the VacMan® Vacuum Manifold and the automated Maxwell RSC instrument into a final volume of 60µL. From all the samples tested, SARS-CoV-2 N1 gene quantification was performed by (RT)-qPCR (probe, primers and cycling conditions described in the CDC-006-00019 CDC/DDID/NCIRD/ Division of Viral Diseases protocol)

Library preparation and probe-based target enrichment sequencing

Prior to library preparation, nucleic acids and a negative control were reverse transcribed into cDNA, which was then tagged and complemented to obtain double-stranded DNA (dsDNA). This randomly tagged dsDNA, representing the viral content, was subsequently amplified through 25 cycles of PCR amplification to generate enough DNA for library preparation, following the methodology described by Fernandez-Cassi *et al.* (Fernandez-Cassi et al., 2018b). Libraries preparation was carried out in duplicate using the KAPA HyperPrep Kit, following the manufacturer's instructions (Roche-Kapa Biosystems, Basel, Switzerland). In brief, the dsDNA was enzymatically fragmented, indexed with Dual Indexes, and amplified. The resulting DNA fragments' quality and concentration were assessed using a Qubit 3.0 Fluorometer (Thermo Fisher Scientific, Massachusetts, USA).

One of the libraries replicates was then subjected to hybridization with probes specifically designed to capture target species from the *Coronaviridae* family utilizing the KAPA HyperCap Enrichment Kit (Roche, Basel, Switzerland), Basel, Switzerland), Basel, Switzerland). Capture probes were designed based on the genomic sequences of 18 species from *Coronaviride* family with the purpose to maximize the hybridization of the probes with these sequences and related ones, and to minimize the capture of any other genomic material that might be present in the environment. Species used for the design were: Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/IND/CD211295/2021, SARS coronavirus Tor2, Middle East respiratory syndrome-related coronavirus isolate HCoV-EMC/2012,

Human coronavirus OC43 strain ATCC VR-759, Human coronavirus 229E, Human Coronavirus NL63, Human coronavirus HKU1, Porcine coronavirus HKU15 strain HKU15-155, Bovine coronavirus isolate BCoV-ENT, Murine hepatitis virus strain A59, Transmissible gastroenteritis virus, Feline infectious peritonitis virus, Avian infectious bronchitis virus, Alphacoronavirus Bat-CoV/P.kuhlii/Italy/3398-19/2015, Bat coronavirus 1A, Rousettus bat coronavirus HKU10, Bat coronavirus HKU8 and Bat coronavirus HKU2 (see Supplementary Information 1).

Following the capture step, the quality and concentration of the captured libraries, non-captured libraries, and negative controls were reassessed. Subsequently, sequencing of these libraries was conducted on an Illumina NextSeq platform (Illumina, California USA), generating an output of 400 million reads.

Bioinformatic analysis

Raw sequencing reads from samples with and without capture—for each set of Sewage (Sewage01, Sewage02 and Sewage03) and Bat Guano (Guano01, Guano02, Guano03 and Guano04) samples, were cleaned, filtered, and assembled using CAPTVRED pipeline ,an automated protocol designed to assess the virome present in complex samples, specially focused on those obtained by capture-based metagenomics approach.The analysis is sensible to degraded genomic sequences that may be present in the environment and provides comprehensive, reproducible and accessible results centered on viral outcomes (Tarradas-Alemany et al, submitted manuscript). It was run locally in a Debian server with 32 threads. Main steps of the pipeline and relevant parameters are briefly described hereafter. For the pipeline cleaning step, the paired FASTQ files were processed with BBDuk (BBMap version 38.96; Institute J.G., 2022), and reads quality was assessed using FastQC (version 0.11.9; Andrews, S., 2010) and MultiQC (version 1.9; Ewels, P. et al, 2016). A second filtering step was performed using Kaiju (version 1.9.0; Menzel P, et al., 2016) to discard reads corresponding to non-viral species (such as eukaryotic, archaea, or bacterial sequences remnant on the samples). After that, paired and unpaired reads were assembled into contigs by Megahit (version 1.0; Li, D. et al., 2016); only those contig sequences and non-assembled reads (singletons) over 100bp were selected. All programs were run using the pipeline default parameters as described in the pipeline documentation. Then, NCBI-BLASTn (version 2.11.0+; Altschul, S. F. et al., 1990) was run using C-

RVDB (Goodacre N. et al, 2018) database as reference (identity > 50% and E-value < 10e-10). Sequenced raw reads obtained from the chicken, porcine and rabbit lixivates samples were fully processed using CAPTVRED complete pipeline which already implemented the BLAST step. The default parameters were used for this analysis.

3. Results and discussion

The use of new targeted methodologies could favor the discovery of potential sources of human infections or the emergence of new zoonotic species or strains. In this study we aimed to validate the efficacy of a TES panel specifically designed for studying members of the *Coronaviridae* family to study the diversity of coronaviruses in sewage and animal samples.

For the samples used, SARS-CoV-2 concentrations ranged from 5,74E+01 GC/ml to 4,10E+03 GC/ml in sewage samples. Guano and animal lixivate samples tested negative.

Panel performance on Coronaviridae family

The abundance of CoV in the studied samples in comparison to the analysis of the same samples by UVM (Untargeted Viral Metagenomics) is graphically represented in Figure 1. In sewage samples, CoV were detected solely when TES was employed. A total of 390,553 reads assigned to diverse CoV species were observed across the three analyzed sewage samples.

Concerning guano samples, sequencing of CoV species was achieved in 2 out of 4 samples. In one of the positive samples, Guano 01, CoV reads were obtained even without using the enrichment panel, resulting in 99,678 reads. However, the implementation of the panel significantly increased the number of reads obtained to 37,537,788, representing an almost three-logarithmic increase.

In the case of animal lixivate samples, only the enrichment strategy enabled the acquisition of CoV reads. Specifically, 58,767 reads were obtained for rabbit lixivates, while pig and chicken lixivate yielded 665 and 300 reads, respectively.

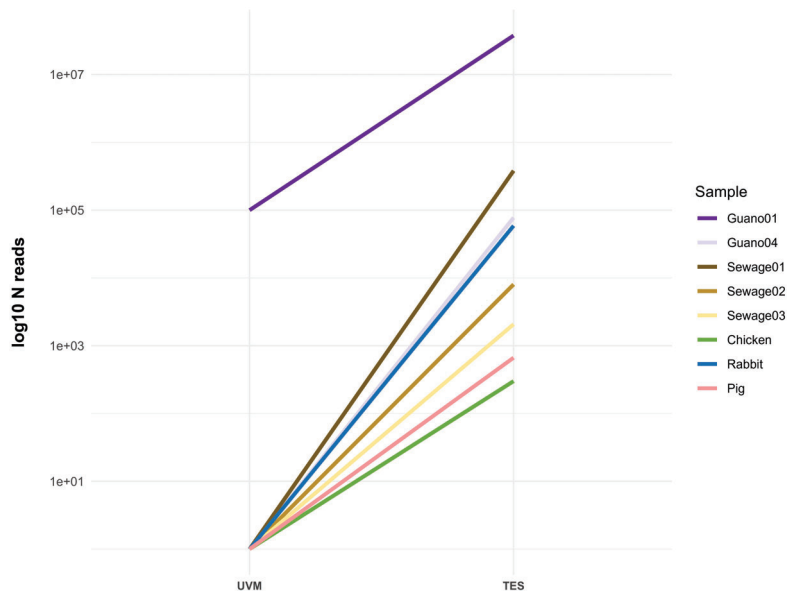


Figure 1. Representation of the Log10 of number of CoV reads showing increase when applying the designed TES approach.

CoV diversity using TES

A wide range of species was detected throughout the samples analyzed when using the enrichment panel, as shown in Figure 2, which presents their abundance and host diversity.

Interestingly, sewage samples exhibited the highest diversity of CoVs. This observation is not surprising, as these complex samples are known to harbor an extensive range of viral diversity (Fernandez-Cassi, et al., 2018b, Martínez-Puchol et al., 2020). Consequently, studying less prevalent species, such as CoVs, within these samples can be challenging and the use of TES methodologies represents the most effective approach to address this limitation.

In this study, the application of a CoV probe panel represented a considerable improvement in reference to other targeted efforts. Previous attempts using a commercial probe-panel directed to vertebrate viruses, including 346 *Coronaviridae* species, only yielded a few reads from SARS-CoV-2, canine CoV, feline CoV, human CoV OC43 and Lucheng Rn rat CoV in urban sewage (Martínez-Puchol et al., 2021).

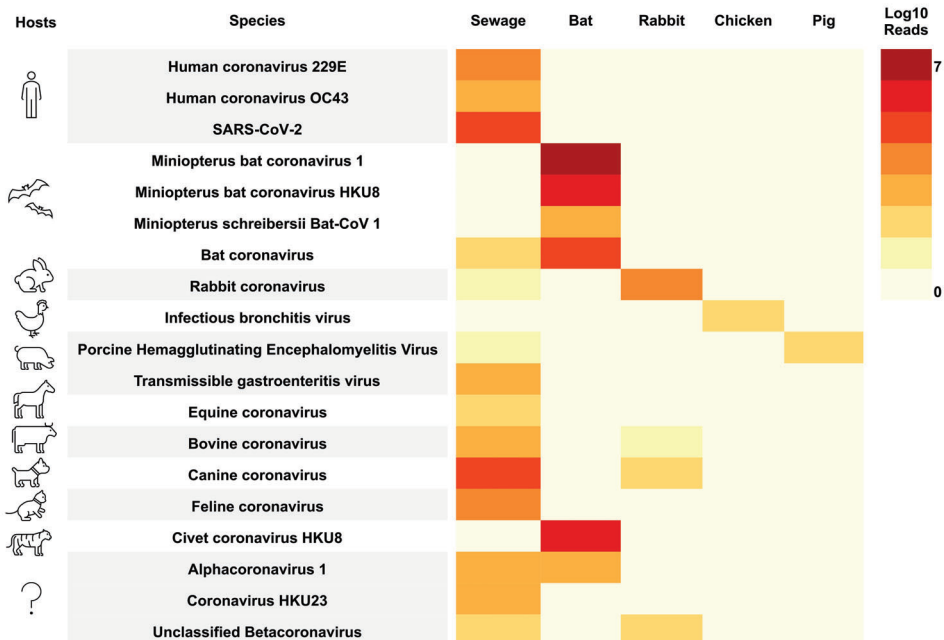


Figure 2. Heatmap representing the CoV species abundance (Log10 number of reads) detected in the analyzed samples.

As shown in Figure 2, SARS-CoV-2 and Canine CoV exhibited the highest number of reads followed by HCoV-229E and Feline CoV. The presence of these species can be attributed to their well-established epidemiological status. SARS-CoV-2 was associated with the ongoing pandemic during the study period (November 2021-January 2022), while HCoV-229E and HCoV OC43 are commonly linked to the occurrence of the common cold (Van der Hoek, 2015). The presence of Canine CoV in the samples can be explained by canine waste disposal through toilets or through streets cleaning which would be applicable to another pet associated CoVs.

The remaining detected species display infectivity across a wide range of animal hosts, providing valuable insights into the contribution of different animal types (domesticated, farm, and wild animals) to the sewage in the studied area. The utilization of these targeted panels not only enhances the understanding of circulating strains but also serves as a potential tool for early detection and response. This highlights the usefulness of this approach in monitoring CoV and other significant

viral families, including certain enteroviruses and poxviruses, as suggested in previous studies (Fernandez-Cassi, et al., 2018a; Hata et al., 2018; Martínez-Puchol et al., 2020).

In the bat guano samples, despite the higher read count compared to other samples, the observed diversity is relatively limited, primarily consisting of four distinct species of CoVs known to infect bats. The more abundant of these species is *Miniopterus bat coronavirus 1*, also known as Bat-CoV MOP1, which belongs to the Alphacoronavirus genus and was the first CoV identified in bats (Poon et al., 2005). Bat-CoV MOP1 is closely related to the other CoV found in the analyzed samples, such as HKU8 and Bat-CoV 1, all of which have a documented impact on *Miniopterus spp.*, commonly known as bent-winged bats (Chu et al., 2008). These bats are considered reservoirs of Alphacoronavirus species (Kimprasit et al., 2021).

In addition to this species, a substantial number of reads associated with Civet CoV HKU8 were also detected in two different guano samples when the enrichment panel was utilized but also without enrichment. Notably, Civet CoV HKU8 was detected in civets, and it was speculated about a potential cross-species transmission event involving a bat-associated CoV (He et al., 2022).

The successful use of the enrichment panel for CoV characterization in bats has been documented in previous studies (Li et al., 2020; Lim et al., 2019). Therefore, the feasibility of conducting targeted, cost-effective, large-scale genome-level surveillance of bat CoVs has been firmly established. The information gained from such surveillance efforts could prove invaluable in the prevention and control of potential spillover events from bats to adjacent mammalian species. Migratory species may contribute to pathogen persistence in species encountered along the migratory path. Given the potential for long-range seasonal movements of *Miniopterus schreibersii* (Serra-Cobo et al. 2009; 2017, Colombi et al. 2019), this specie may represent a central vector for spatial dispersion of viruses in southern Europe, where this bat species is abundant. Determining which bat colonies exhibit higher prevalence of viruses with zoonotic potential enables the implementation of preventive measures to mitigate the transmission to the human population, livestock, and pets.

In the analysis of animal lixivate samples, a lower viral diversity was detected in comparison to the results obtained from wastewater and guano samples. This observation suggests that these highly complex samples may have a lower contribution of viral excretion, a higher degree of viral nucleic acid degradation, or may require an optimized methodology for viral isolation.

266 Infectious bronchitis virus (IBV) reads were detected in chicken lixiviates. IBV is a particularly
267 relevant virus that infects various avian body compartments, including the respiratory tract, kidney,
268 gut, and reproductive systems, and is a significant contributor to economic losses within the poultry
269 industry (Cavanagh, 2007). Since live attenuated vaccine is available against IBV, the possibility that
270 some of these sequences belong to excreted vaccine strains could not be ruled out.

271 The continuous emergence of novel viral strains through mutations and recombination events in the
272 viral genome exacerbates the challenges associated with the identification and control of IBV
273 (Jackwood, 2012). Given its status as a major respiratory pathogen affecting the European poultry
274 industry, advancements in IBV detection, particularly in conjunction with other relevant CoVs,
275 utilizing enrichment panels, could hold the key to overcoming the apparent difficulties in its control.

276 Porcine hemagglutinating encephalomyelitis virus (PHEV) sequences were found in pig lixiviates
277 samples. PHEV is responsible for causing vomiting and/or encephalomyelitis in pigs, being the sole
278 identified neurotropic CoV known to affect pigs. PHEV exhibits a high prevalence and commonly
279 circulates sub clinically within most swine herds globally, not having significant clinical implications
280 in many swine-producing countries (Mora-Díaz et al., 2019). While PHEV is not currently recognized
281 as a zoonotic virus, it is crucial to acknowledge the emerging understanding of cross-species
282 transmission and zoonotic potential demonstrated by other porcine CoVs such as Porcine delta CoV
283 and Swine acute diarrhea syndrome CoV (Guo et al., 2023). This knowledge underscores the
284 importance of implementing effective control measures for CoVs present in pigs.

285 In the sequencing analysis of rabbit lixiviates, the presence of Rabbit CoV, Bovine CoV, and Canine
286 CoV was observed. Rabbit CoV is recognized as the causative agent of common enteric infections in
287 rabbit colonies (Descôteaux et al., 1985). However, the extent of its involvement in cross-species
288 transmission events, as well as the implications of Bovine CoV, remain largely unknown. Canine
289 CoV, which was also detected in sewage samples, has been associated with febrile or lower
290 respiratory symptoms in humans (Lednicky et al., 2022; Vlasova et al., 2022), suggesting possibilities
291 for cross-species transmission. Therefore, controlling its spread is significant to prevent potential
292 recombinant transmissions among various mammalian species.

293 In addition to identifying the obtained species from each sample, further characterization of these
294 viral assignments was conducted based on genomic coverage and identity, as outlined in Table 1.

Table 1. CoV assignments obtained by TES and UVM from analyzed samples including genome coverage, and nucleotide identity represented in percentages.

Sample	Coronaviridae species	TES			UVM		
		Reads (log10 N)	Genome coverage (%)	Identity (%)	Reads (log10 N)	Genome coverage (%)	Identity (%)
Sewage	Alphacoronavirus 1	3,45	96,50	99,87	ND	-	-
	Bat coronavirus	2,97	99,97	97,89	ND	-	-
	Bovine coronavirus	3,94	4,64	97,29	ND	-	-
	Canine coronavirus	2,96	95,65	92,60	ND	-	-
	Coronavirus HKU23	3,95	73,08	98,07	ND	-	-
	Equine coronavirus	2,43	71,04	93,69	ND	-	-
	Feline coronavirus	3,97	92,30	83,87	ND	-	-
	Human coronavirus 229E	4,34	85,59	99,17	ND	-	-
	Human coronavirus OC43	3,93	74,36	95,33	ND	-	-
	Porcine hemagglutinating encephalomyelitis virus	1,76	73,43	95,22	ND	-	-
	Rabbit coronavirus	1,32	71,60	93,44	ND	-	-
	SARS-CoV-2	5,80	99,96	99,89	ND	-	-
	Transmissible gastroenteritis virus	5,80	97,18	92,24	ND	-	-
	Unclassified Betacoronavirus	1,89	99,51	90,98	ND	-	-
Bat guano	Alphacoronavirus 1	3,94	73,01	74,45	ND	-	-
	Bat coronavirus	5,57	99,96	79,81	2,77	91,02	79,60
	Civet coronavirus HKU8	6,14	85,86	96,63	2,31	82,22	96,37
	Miniopterus bat coronavirus 1	7,54	98,45	97,25	4,99	98,27	96,06
	Miniopterus bat coronavirus HKU8	6,09	98,33	86,11	2,40	88,56	84,69
	Miniopterus schreibersii bat coronavirus 1	3,79	93,67	74,68	ND	-	-
Chicken	Infectious bronchitis virus	2,48	1,71	100,00	ND	-	-
Rabbit	Rabbit coronavirus	4,76	2,85	97,74	ND	-	-
	Canine coronavirus	2,94	2,43	96,43	ND	-	-
	Bovine coronavirus	1,86	0,75	98,21	ND	-	-
	Unclassified Betacoronavirus	2,49	0,73	99,04	ND	-	-
Pig	Porcine hemagglutinating encephalomyelitis virus	2,82	1,16	98,42	ND	-	-

The use of TES led to notable differences in read acquisition, especially in sewage and bat samples, where substantial genome coverage was achieved for viral assignments when employing the enrichment panel. Conversely, in the unenriched sample (UVM) where reads were obtained without the enrichment, a lower level of genomic coverage was observed, consistent with previous comparisons of these NGS techniques (Martínez-Puchol et al., 2021).

The coverage observed in the animal lixivate samples was relatively low, potentially attributed to the lower quantity of reads obtained for certain species that could have been caused, as explained before, for the lower contribution of viral excretion in these samples or for the viral nucleic acid degradation.

4. Conclusions

The probe-based enrichment panel designed for sequencing members of the *Coronaviridae* family demonstrated success in capturing and sequencing not only the targeted species but also other viral species within the family. A higher number of different CoV species were detected in comparison with untargeted metagenomics or with vertebrate viruses directed panels applied in other studies by this research group. This enrichment panel provided increased sequencing depth for the intended viral species. It also showed potential for viral discovery which could enable in the future comprehensive studies on both known and potential animal species susceptible to spillover events. This approach holds promise for enhancing our understanding of viral dynamics and their potential implications for zoonotic transmission.

Acknowledgments

The authors express their gratitude to the staff of the rendering and wastewater treatment plants involved in this study for their assistance and technical support during sampling.

The authors wish to express their gratitude to Ferran Briansó from Roche Diagnostics for the collaborative efforts in developing the capture probe panel.

This research was conducted as part of the EPISARS project, which received funding from donations obtained during “La Marató de TV3” dedicated to COVID-19 and was also funded by MCIN/AEI/10.13039/501100011033 and by “ERDF A way of making Europe”.

S. M-P holds a *Margarita Salas* grant funded by Ministerio de Universidades - NextGeneration EU.

S. B-M is a Serra-Hunter fellow at the University of Barcelona.

Availability of data and material

Raw sequencing data is available from:

https://compgen.bio.ub.edu/datasets/Target_Enrichment_MetaViromics_CoV/rawdata/

During the preparation of this work the authors used <https://chat.openai.com/> to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, 215(3), 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. Babraham Institute. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

Briese, T., Kapoor, A., Mishra, N., Jain, K., Kumar, A., Jabado, O. J., & Ian Lipkina, W. (2015). Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *MBio*, 6(5), 1–12. <https://doi.org/10.1128/mBio.01491-15>

Cantalupo, P. G., Calgua, B., Zhao, G., Hundesa, A., Wier, A. D., Katz, J. P., Grabe, M., Hendrix, R. W., Girones, R., Wang, D., & Pipas, J. M. (2011). Raw sewage harbors diverse viral populations. *MBio*, 2(5), e00180-11. <https://doi.org/10.1128/mBio.00180-11>

Carlson, C. J., Gibb, R. J., Albery, G. F., Brierley, L., Connor, R. P., Dallas, T. A., Eskew, E. A., Fagre, A. C., Farrell, M. J., Frank, H. K., Muylaert, R. L., Poisot, T., Rasmussen, A. L., Ryan, S. J., & Seifert, S. N. (2022). The Global Virome in One Network (VIRION): an Atlas of Vertebrate-Virus Associations. *MBio*, 13(2). <https://doi.org/10.1128/mbio.02985-21>

Cavanagh, D. (2007). Coronavirus avian infectious bronchitis virus. In *Veterinary Research* (Vol. 38, Issue 2). <https://doi.org/10.1051/vetres:2006055>

353 Charre, C., Ginevra, C., Sabatier, M., Regue, H., Destras, G., Brun, S., Burfin, G., Scholtes, C.,
354 Morfin, F., Valette, M., Lina, B., Bal, A., & Josset, L. (2020). Evaluation of NGS-based
355 approaches for SARS-CoV-2 whole genome characterisation. *Virus Evolution*.
356 <https://doi.org/10.1093/ve/veaa075>

357 Chu, D. K. W., Peiris, J. S. M., Chen, H., Guan, Y., & Poon, L. L. M. (2008). Genomic
358 characterizations of bat coronaviruses (1A, 1B and HKU8) and evidence for co-infections in
359 *Miniopterus* bats. *Journal of General Virology*, 89(5). <https://doi.org/10.1099/vir.0.83605-0>

360 Colombi, D., Serra-Cobo, J., Métras, R., Apolloni, A., Poletto, Ch., López-Roig, M., Bourhy, H., &
361 Colizza, V. (2019). Mechanisms for lyssavirus persistence in non-synanthropic bats in Europe:
362 insights from a modeling study. *Scientific Reports*, 9, 537 DOI:10.1038/s41598-018-36485-y

363 Descôteaux, J. P., Lussier, G., Berthiaume, L., Alain, R., Seguin, C., & Trudel, M. (1985). An enteric
364 coronavirus of the rabbit: Detection by immunoelectron microscopy and identification of
365 structural polypeptides. *Archives of Virology*, 84(3–4). <https://doi.org/10.1007/BF01378976>

366 Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for
367 multiple tools and samples in a single report. *Bioinformatics* (Oxford, England), 32(19), 3047–
368 3048. <https://doi.org/10.1093/bioinformatics/btw354>

369 Fernandez-Cassi, X., Rusiñol, M., & Martínez-Puchol, S. (2018a). Viral concentration and
370 amplification from human serum samples prior to application of next-generation sequencing
371 analysis. In *Methods in Molecular Biology* (Vol. 1838, pp. 173–188).
372 https://doi.org/10.1007/978-1-4939-8682-8_13

373 Fernandez-Cassi, X., Timoneda, N., Martínez-Puchol, S., Rusiñol, M., Rodríguez-Manzano, J.,
374 Figuerola, N., Bofill-Mas, S., Abril, J. F., & Girones, R. (2018b). Metagenomics for the study of
375 viruses in urban sewage as a tool for public health surveillance. *Science of The Total*
376 *Environment*, 618(August), 870–880. <https://doi.org/10.1016/j.scitotenv.2017.08.249>

377 Forés, E., Bofill-Mas, S., Itarte, M., Martínez-Puchol, S., Hundesa, A., Calvo, M., Borrego, C. M.,
378 Corominas, L. L., Girones, R., & Rusiñol, M. (2021). Evaluation of two rapid ultrafiltration-
379 based methods for SARS-CoV-2 concentration from wastewater. *Science of the Total*
380 *Environment*, 768. <https://doi.org/10.1016/j.scitotenv.2020.144786>

381 Goodacre, N., Aljanahi, A., Nandakumar, S., Mikailov, M., & Khan, A. S. (2018). A Reference Viral
382 Database (RVDB) To Enhance Bioinformatics Analysis of High-Throughput Sequencing for
383 Novel Virus Detection. *mSphere*, 3(2), e00069-
384 18. <https://doi.org/10.1128/mSphereDirect.00069-18>

385 Guo, Z., Jin, Q., Li, P., Xing, G., Lu, Q., & Zhang, G. (2023). Potential cross-species transmission
386 risks of emerging swine enteric coronavirus to human beings. *Journal of Medical Virology*,
387 95(7). <https://doi.org/10.1002/jmv.28919>

388 Hata, A., Kitajima, M., Haramoto, E., Lee, S., Ihara, M., Gerba, C. P., & Tanaka, H. (2018). Next-
389 generation amplicon sequencing identifies genetically diverse human astroviruses, including
390 recombinant strains, in environmental waters. *Scientific Reports*, 8(1), 11837.
391 <https://doi.org/10.1038/s41598-018-30217-y>

392 He, W. T., Hou, X., Zhao, J., Sun, J., He, H., Si, W., Wang, J., Jiang, Z., Yan, Z., Xing, G., Lu, M.,
393 Suchard, M. A., Ji, X., Gong, W., He, B., Li, J., Lemey, P., Guo, D., Tu, C., ... Su, S. (2022).
394 Virome characterization of game animals in China reveals a spectrum of emerging pathogens.
395 *Cell*, 185(7). <https://doi.org/10.1016/j.cell.2022.02.014>

396 Institute, J. G. (2022). BBDuk, BBMap Tools. <https://shorturl.at/qJOS4>

397 Islam, A., Ferdous, J., Islam, S., Sayeed, M. A., Rahman, M. K., Saha, O., Hassan, M. M., & Shirin,
398 T. (2022). Transmission dynamics and susceptibility patterns of SARS-CoV-2 in domestic,
399 farmed and wild animals: Sustainable One Health surveillance for conservation and public health
400 to prevent future epidemics and pandemics. In *Transboundary and Emerging Diseases* (Vol. 69,
401 Issue 5). <https://doi.org/10.1111/tbed.14356>

402 Jackwood, M. W. (2012). Review of infectious bronchitis virus around the world. In *Avian Diseases*
403 (Vol. 56, Issue 4). <https://doi.org/10.1637/10227-043012-Review.1>

404 Jo, W. K., Drosten, C., & Drexler, J. F. (2021). The evolutionary dynamics of endemic human
405 coronaviruses. *Virus Evolution*, 7(1). <https://doi.org/10.1093/ve/veab020>

406 Khamassi Khbou, M., Daaloul Jedidi, M., Bouaicha Zaafour, F., & Benzarti, M. (2021).
407 Coronaviruses in farm animals: Epidemiology and public health implications. In *Veterinary*
408 *Medicine and Science* (Vol. 7, Issue 2). <https://doi.org/10.1002/vms3.359>

409 Kimprasit, T., Nunome, M., Iida, K., Murakami, Y., Wong, M. L., Wu, C. H., Kobayashi, R.,
410 Hengjan, Y., Takemae, H., Yonemitsu, K., Kuwata, R., Shimoda, H., Si, L., Sohn, J. H.,
411 Asakawa, S., Ichiyanagi, K., Maeda, K., Oh, H. S., Mizutani, T., ... Hondo, E. (2021). Dispersal
412 history of *Miniopterus fuliginosus* bats and their associated viruses in east Asia. *PLoS ONE*, 16(1
413 January). <https://doi.org/10.1371/journal.pone.0244006>

414 Lednicky, J. A., Tagliamonte, M. S., White, S. K., Blohm, G. M., Alam, M. M., Iovine, N. M., Salemi,
415 M., Mavian, C., & Morris, J. G. (2022). Isolation of a Novel Recombinant Canine Coronavirus
416 from a Visitor to Haiti: Further Evidence of Transmission of Coronaviruses of Zoonotic Origin
417 to Humans. *Clinical Infectious Diseases*, 75(1). <https://doi.org/10.1093/cid/ciab924>

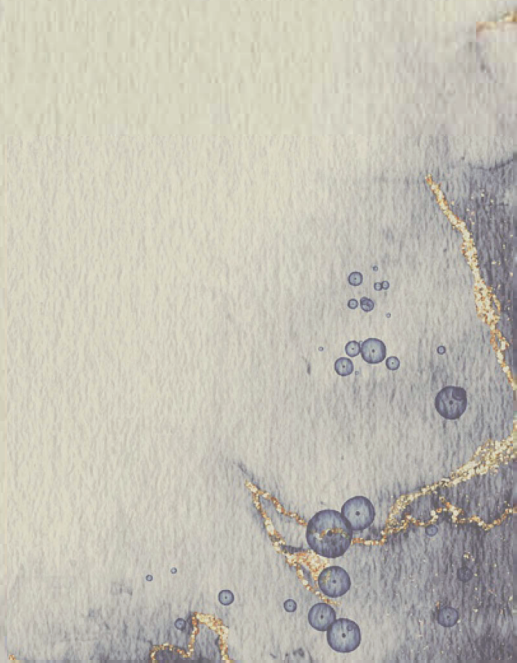
418 Li, D., Luo, R., Liu, C. M., Leung, C. M., Ting, H. F., Sadakane, K., Yamashita, H., & Lam, T. W.
419 (2016). MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced
420 methodologies and community practices. *Methods (San Diego, Calif.)*, 102, 3–
421 11. <https://doi.org/10.1016/j.ymeth.2016.02.020>

- Li, B., Si, H.-R., Zhu, Y., Yang, X.-L., Anderson, D. E., Shi, Z.-L., Wang, L.-F., & Zhou, P. (2020).
Discovery of Bat Coronaviruses through Surveillance and Probe Capture-Based Next-
Generation Sequencing. *MSphere*, 5(1), 1–10. <https://doi.org/10.1128/msphere.00807-19>
- Lim, X. F., Lee, C. B., Pascoe, S. M., How, C. B., Chan, S., Tan, J. H., Yang, X., Zhou, P., Shi, Z.,
Sessions, O. M., Wang, L. F., Ng, L. C., Anderson, D. E., & Yap, G. (2019). Detection and
characterization of a novel bat-borne coronavirus in Singapore using multiple molecular
approaches. *The Journal of General Virology*, 100(10), 1363–1374.
<https://doi.org/10.1099/jgv.0.001307>
- Liu, H., Li, J., Lin, Y., Bo, X., Song, H., Li, K., Li, P., & Ni, M. (2022). Assessment of two-pool
multiplex long-amplicon nanopore sequencing of SARS-CoV-2. *Journal of Medical Virology*,
94(1), 327–334. <https://doi.org/10.1002/jmv.27336>
- Martínez-Puchol, S., Itarte, M., Rusiñol, M., Forés, E., Mejías-Molina, C., Andrés, C., Antón, A.,
Quer, J., Abril, J. F., Girones, R., & Bofill-Mas, S. (2021). Exploring the diversity of coronavirus
in sewage during COVID-19 pandemic: Don't miss the forest for the trees. *Science of the Total
Environment*, 800. <https://doi.org/10.1016/j.scitotenv.2021.149562>
- Martínez-Puchol, S., Rusiñol, M., Fernández-Cassi, X., Timoneda, N., Itarte, M., Andrés, C., Antón,
A., Abril, J. F., Girones, R., & Bofill-Mas, S. (2020). Characterisation of the sewage virome:
comparison of NGS tools and occurrence of significant pathogens. *Science of the Total
Environment*, 713. <https://doi.org/10.1016/j.scitotenv.2020.136604>
- Maryam S, Ul Haq I, Yahya G, Ul Haq M, Alghammal AM, Saber S, Cavalu S. (2023). COVID-19
surveillance in wastewater: An epidemiological tool for the monitoring of SARS-CoV-2. *Front
Cell Infect Microbiol*, 12:978643. doi: 10.3389/fcimb.2022.978643.
- Menzel, P., Ng, K. L., & Krogh, A. (2016). Fast and sensitive taxonomic classification for
metagenomics with Kaiju. *Nature communications*, 7,
11257. <https://doi.org/10.1038/ncomms11257>
- Mora-Díaz, J. C., Piñeyro, P. E., Houston, E., Zimmerman, J., & Giménez-Lirola, L. G. (2019).
Porcine hemagglutinating encephalomyelitis virus: A review. In *Frontiers in Veterinary Science*
(Vol. 6, Issue FEB). <https://doi.org/10.3389/fvets.2019.00053>
- Paskey, A. C., Frey, K. G., Schroth, G., Gross, S., Hamilton, T., & Bishop-Lilly, K. A. (2019).
Enrichment post-library preparation enhances the sensitivity of high-throughput sequencing-
based detection and characterization of viruses from complex samples. *BMC Genomics*, 20(1),
155. <https://doi.org/10.1186/s12864-019-5543-2>

- Pekar, J. E., Magee, A., Parker, E., Moshiri, N., Izhikevich, K., Havens, J. L., Gangavarapu, K., Serrano, L. M. M., Crits-Christoph, A., Matteson, N. L., Zeller, M., Levy, J. I., Wang, J. C., Hughes, S., Lee, J., Park, H., Park, M. S., Yan, K. C. Z., Lin, R. T. P., ... Wertheim, J. O. (2022). The molecular epidemiology of multiple zoonotic origins of SARS-CoV-2. *Science*, 377(6609). <https://doi.org/10.1126/science.abp8337>
- Pina, S., Puig, M., Lucena, F., Jofre, J., & Girones, R. (1998). Viral pollution in the environment and in shellfish: Human adenovirus detection by PCR as an index of human viruses. *Applied and Environmental Microbiology*, 64(9). <https://doi.org/10.1128/aem.64.9.3376-3382.1998>
- Poon, L. L. M., Chu, D. K. W., Chan, K. H., Wong, O. K., Ellis, T. M., Leung, Y. H. C., Lau, S. K. P., Woo, P. C. Y., Suen, K. Y., Yuen, K. Y., Guan, Y., & Peiris, J. S. M. (2005). Identification of a Novel Coronavirus in Bats. *Journal of Virology*, 79(4), 2001–2009. <https://doi.org/10.1128/jvi.79.4.2001-2009.2005>
- Ruiz-Aravena, M., McKee, C., Gamble, A., Lunn, T., Morris, A., Snedden, C. E., Yinda, C. K., Port, J. R., Buchholz, D. W., Yeo, Y. Y., Faust, C., Jax, E., Dee, L., Jones, D. N., Kessler, M. K., Falvo, C., Crowley, D., Bharti, N., Brook, C. E., ... Plowright, R. K. (2022). Ecology, evolution and spillover of coronaviruses from bats. In *Nature Reviews Microbiology* (Vol. 20, Issue 5). <https://doi.org/10.1038/s41579-021-00652-2>
- Serra-Cobo, J., Sanz-Trullén, V. & Martínez-Rica, J. P. Migratory movements of *Miniopterus schreibersii* in the north-east of Spain. (2017). *Acta Theriol. (Warsz.)* 43, 271–283
- Serra-Cobo, J., López-Roig, M., Bayer, X., Amengual, B., Guasch C. (2009). *Bats: Science and Myth*. Barcelona, Publications and Editions, University of Barcelona; 268 pages.
- Sharun, K., Dhama, K., Pawde, A. M., Gortázar, C., Tiwari, R., Bonilla-Aldana, D. K., Rodríguez-Morales, A. J., de la Fuente, J., Michalak, I., & Attia, Y. A. (2021). SARS-CoV-2 in animals: potential for unknown reservoir hosts and public health implications. In *Veterinary Quarterly* (Vol. 41, Issue 1). <https://doi.org/10.1080/01652176.2021.1921311>
- Van der Hoek, L. (2015). Human coronaviruses: What do they cause? *Antiviral Therapy*, 12(March).
- Vlasova, A. N., Diaz, A., Damtie, D., Xiu, L., Toh, T. H., Lee, J. S. Y., Saif, L. J., & Gray, G. C. (2022). Novel Canine Coronavirus Isolated from a Hospitalized Patient with Pneumonia in East Malaysia. *Clinical Infectious Diseases*, 74(3). <https://doi.org/10.1093/cid/ciab456>
- Zhou, H., Chen, X., Hu, T., Li, J., Song, H., Liu, Y., Wang, P., Liu, D., Yang, J., Holmes, E. C., Hughes, A. C., Bi, Y., & Shi, W. (2020). A Novel Bat Coronavirus Closely Related to SARS-CoV-2 Contains Natural Insertions at the S1/S2 Cleavage Site of the Spike Protein. *Current Biology*, 30(11), 2196-2203.e3. <https://doi.org/10.1016/j.cub.2020.05.023>

Capítol 2

Exposició a virus contaminants
d'aigua i aliments:
detecció i caracterització vírica.



Els virus excretats per humans i animals que arriben a les aigües residuals poden ser resistents als tractaments de les estacions depuradores d'aigües residuals i acabar contaminant les fonts d'aigua que s'utilitzen per beure, per regar o per a ús recreatiu, constituint així potencials vies de transmissió d'aquests virus. En aquest capítol s'estudia l'exposició a virus patògens presents en aigües i aliments mitjançant la detecció i caracterització vírica, especialment aplicant tècniques de NGS en aliments de producció orgànica i les seves fonts d'aigua de reg. A més, també s'avalua la contaminació vírica i s'explora la diversitat de virus presents en esorrenties i aigües subterrànies d'un entorn urbà, les quals, en l'actual context de sequera, són de gran interès per a diferents usos.

Article 4: Les tècniques de NGS revelen una gran diversitat de virus RNA i papil·lomavirus en aliments de producció orgànica i en aigües de reg

“NGS Techniques Reveal a High Diversity of RNA Viral Pathogens and Papillomaviruses in Fresh Produce and Irrigation Water”

Itarte, M., Martínez-Puchol, S., Forés, E., Hundesa, A., Timoneda, N., Bofill-Mas, S., Girones, R. i Rusiñol, M.

Foods 2021. Volume 10 (8), 1820. <https://doi.org/10.3390/foods10081820>

Les fruites i verdures són susceptibles a la contaminació microbiològica en cada etapa de la cadena de producció alimentària i, com a font potencial de patògens, la qualitat de l'aigua de reg esdevé un factor crític. En els últims anys, les tècniques de seqüenciació massiva han anat prosperant i s'han estès a una àmplia varietat de camps. No obstant això, la seva aplicació en la seguretat alimentària encara està poc explorada, i es necessiten millores en la seva sensibilitat.

En aquest estudi es va detectar, mitjançant assaigs de qPCR, una contaminació baixa però freqüent de patògens vírics que circulen comunament en el 46.9% de les mostres d'aliments de producció orgànica analitzades: 6/12 mostres d'enciam, 4/12 mostres de maduixes i 5/8 mostres de julivert. A més, l'aplicació de dues tècniques de NGS, la seqüenciació massiva amb enriquiment de dianes mitjançant el panell de captura VirCapSeq-VERT, per a la detecció de virus que infecten vertebrats, i la seqüenciació massiva d'amplicons, va revelar una alta diversitat de patògens vírics, especialment NoV i HPV, en fruita, verdura i aigua de reg. Tots els tipus de NoV i HPV identificats en les mostres de fruita i verdura també es van detectar en diverses aigües de reg, indicant

que aquests virus patògens circulen i són excretats per la població, i que l'aigua de reg és la font més probable dels patògens vírics identificats en aquestes mostres d'aliments de producció orgànica.

Les tres metodologies utilitzades per a l'anàlisi de la contaminació vírica en mostres d'aliments i mostres d'aigua de reg van demostrar ser útils i van proporcionar diferents tipus d'informació a l'estudi: (a) la qPCR és una tècnica quantitativa, sensible i específica, (b) la seqüenciació massiva amb captura utilitzant el panell VirCapSeq-VERT permet la detecció de patògens vírics, relativament abundants, presents en l'aigua de reg, incloent-hi patògens vírics inesperats i soques potencialment zoonòtiques, i (c) la seqüenciació massiva d'amplicons proporciona una sensibilitat més alta per a la identificació de tipus o variants víriques que contaminen aigües i aliments dins d'un grup determinat de virus.

Article

NGS Techniques Reveal a High Diversity of RNA Viral Pathogens and Papillomaviruses in Fresh Produce and Irrigation Water

Marta Itarte ^{1,2}, Sandra Martínez-Puchol ^{1,2}, Eva Forés ^{1,2}, Ayalkibet Hundesa ^{1,2}, Natàlia Timoneda ³,
Sílvia Bofill-Mas ^{1,2}, Rosina Girones ^{1,2,*} and Marta Rusiñol ⁴

- ¹ Laboratory of Viruses Contaminants of Water and Food, Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, 08028 Barcelona, Spain; mitarte@ub.edu (M.I.); smartinezpuchol@ub.edu (S.M.-P.); efores@ub.edu (E.F.); ahundesa@ub.edu (A.H.); sbofill@ub.edu (S.B.-M.)
 - ² The Water Research Institute (IdRA), University of Barcelona, 08001 Barcelona, Spain
 - ³ Department of Marine Biology and Oceanography, Institute of Marine Science, Consejo Superior de Investigaciones Científicas (CSIC), 08003 Barcelona, Spain; timoneda@icm.csic.es
 - ⁴ Institute of Environmental Assessment & Water Research (IDAEA), Consejo Superior de Investigaciones Científicas (CSIC), 08034 Barcelona, Spain; marta.rosinol@idaea.csic.es
- * Correspondence: rgirones@ub.edu

Abstract: Fresh fruits and vegetables are susceptible to microbial contamination at every stage of the food production chain, and as a potential source of pathogens, irrigation water quality is a critical factor. Next-generation sequencing (NGS) techniques have been flourishing and expanding to a wide variety of fields. However, their application in food safety remains insufficiently explored, and their sensitivity requires improvement. In this study, quantitative polymerase chain reaction (qPCR) assays showed low but frequent contamination of common circulating viral pathogens, which were found in 46.9% of samples of fresh produce: 6/12 lettuce samples, 4/12 strawberries samples, and 5/8 parsley samples. Furthermore, the application of two different NGS approaches, target enrichment sequencing (TES) for detecting viruses that infect vertebrates and amplicon deep sequencing (ADS), revealed a high diversity of viral pathogens, especially Norovirus (NoV) and Human Papillomavirus (HPV), in fresh produce and irrigation water. All NoV and HPV types found in fresh fruit and vegetable samples were also detected in irrigation water sources, indicating that these viruses are common circulating pathogens in the population and that irrigation water may be the most probable source of viral pathogens in food samples.

Keywords: organic food; irrigation water; viral pathogens; food safety; next-generation sequencing; target enrichment sequencing; amplicon deep sequencing; human papillomavirus; norovirus; vertebrate viruses



Citation: Itarte, M.; Martínez-Puchol, S.; Forés, E.; Hundesa, A.; Timoneda, N.; Bofill-Mas, S.; Girones, R.; Rusiñol, M. NGS Techniques Reveal a High Diversity of RNA Viral Pathogens and Papillomaviruses in Fresh Produce and Irrigation Water. *Foods* **2021**, *10*, 1820. <https://doi.org/10.3390/foods10081820>

Academic Editor: Dario De Medici

Received: 4 May 2021

Accepted: 31 July 2021

Published: 6 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Consumption and production of fresh fruits and vegetables have increased over the last few years [1] due to population growth, changes in human lifestyles and growing awareness of the benefits of these foods as important sources of nutritional compounds in a healthy and balanced diet [2–4]. This increase in fresh food consumption, often eaten raw or minimally processed, has also been associated with an increase in foodborne infections and disease outbreaks, most of which have been linked to viral origins [2,5,6]. Pathogen contamination can occur at any stage of the food production process, from farm to fork, and irrigation water quality is a critical factor, since it is a potential source of foodborne pathogens [7–10], especially if it comes in direct contact with the edible portion [11]. The microbial quality of irrigation water is affected by a wide range of agricultural, wildlife, and human factors, including: growing season, geographical location, land use, surrounding activities, and environmental conditions [12–14]. Agricultural farms obtain irrigation

water from sources such as reservoirs, rivers and groundwaters, and in the context of the circular economy, reclaimed water is increasingly becoming an important water source for irrigation.

In 2019, Machado-Moreira et al. collected data from several publications and established that most reported outbreaks related to ‘ready-to-eat’ food between 1980 and 2016 were due to the consumption of leafy green vegetables, including lettuce and parsley, whereas strawberries ranked first place when considering the number of cases linked to the consumption of a particular foodstuff [15]. This review also reported that leafy green vegetables and soft fruits are, in fact, the main foodstuffs implicated in the transmission of Norovirus (NoV). This virus, excreted in human feces and transmitted by the fecal–oral route, is the leading cause of reported foodborne disease outbreaks and is the main cause of viral gastroenteritis in people of all ages worldwide [5,16–19]. Other enteric viruses, such as Human Adenovirus (HAdV), Rotavirus (RoV), Hepatitis A Virus (HAV), Hepatitis E Virus (HEV), and Astrovirus (Ast), are also important agents in foodborne outbreaks [7,14,20]. Most pathogenic microorganisms of fecal origin that may be present in irrigation water cause gastroenteritis or acute hepatitis, but other pathologies—such as meningitis, myocarditis, and neurological disorders—are also possible [5].

Human papillomavirus (HPV) infects the skin and mucosal epithelia, with effects ranging from benign lesions, such as common warts, to malignant carcinomas, and its occurrence has been described in raw sewage and river waters [21–24], including high- and low-risk oncogenic HPV types [21,24]. In fact, HPV excretion in the feces of patients with diarrhea has been described, suggesting transmission through fecal shedding of a virus that was believed to be mainly epitheliotropic and, therefore, pointing to possible transmission through contaminated water [23,25].

In most foodborne viral outbreaks, the link between the contaminated food consumed and the people infected is often not easily established, which makes it difficult to intervene and implement preventive measures [26]. Current guidelines on microbiological irrigation water quality and safety from the European Union rely only on the use of *Escherichia coli* (EC) as Fecal Indicator Bacteria (FIB) [27,28]. However, it is well established that these indicators do not always correlate with important waterborne pathogens that may be present in diverse water sources, such as irrigation water [5,29,30]. This is particularly relevant for viral pathogens, which are more resistant to water treatments than bacteria, and thus, FIB might not accurately represent viral inactivation [31–33]. A more suitable indicator for viral fecal contamination is HAdV since it is shed in high concentrations and does not show seasonal variability [34]. HAdV is widely used as a viral fecal indicator; it is highly stable under many environmental conditions and disinfection treatments [29,35–37].

Quantitative polymerase chain reaction (qPCR) is the most commonly used method for the quantification of viruses in food [26], and standardized assays for NoV and HAV (ISO 1526-1:2017, <https://www.iso.org/standard/65681.html>, accessed on 8 April 2021) have been established [38]. However, the detection of viruses in some food matrices can be difficult due to the presence of inhibitory substances that may impact qPCR detection, leading to false-negative results [39–42]. Next-generation sequencing (NGS) is a promising tool with clear applicability to the detection of viral pathogens in the field of food safety [5,43–45]. The introduction of NGS techniques to this field allows the simultaneous analysis of myriad viral pathogens in a single assay, including pathogens initially not suspected to be present. This is achieved by identifying viral sequences in a sample and comparing them with established sequences in databases [39]. Therefore, NGS has great potential as a viral surveillance tool in the food production chain due to its sensitivity, broad detection range, and detailed information about the detected virus [46,47]. Despite this potential, NGS has still not been widely explored in food safety studies, and technical optimization is needed since there are limitations associated with the presence of inhibitory substances [5]. Several viral metagenomic studies applying NGS techniques in the context of food safety have been published [45,48–53], but to our knowledge, only a few have

focused on fresh produce: specifically, frozen berries [54], strawberries [39], lettuce [43], parsley [5], and celery [55].

In this study, lettuce, strawberry, and parsley samples grown under organic agricultural practices as well as water irrigation sources were analyzed by qPCR for the presence and concentration of HAdV, which was used as a fecal viral indicator, and other relevant pathogens, such as NoV GI, NoV GII, and HEV. In addition, two different NGS approaches, target enrichment sequencing (TES) for the analysis of viruses that infect vertebrates and amplicon deep sequencing (ADS) for NoV and HPV, were applied to the studied samples to explore the potential application of NGS techniques for viral detection, characterization, and discovery, especially in fresh fruits and vegetables.

2. Materials and Methods

2.1. Organic Food and Water Samples

Lettuce ($n = 12$), strawberry ($n = 12$) and parsley ($n = 8$) samples were directly purchased from three different organic agriculture producers located in the province of Barcelona (Production Sites 1, 2, and 3 in Table 1), all of which are regulated by Commission Regulation (EC) No. 834/2007, No. 889/2008, and No. 1235/2008 and authorized by Consell Català de la Producció Agrària Ecològica (CCPAE). The irrigation water sources at these production sites were also sampled: 50 L of groundwater samples ($n = 8$) and 50 L of river water (flow of 0.48 m³/s) samples ($n = 4$). Fresh produce and irrigation water samples from organic agriculture producers were collected in July 2019.

Table 1. Irrigation water (GW: groundwater; RIV: river), lettuce (LET), strawberry (STR), and parsley (PAR) samples collected from Production Sites 1, 2, and 3 and also alternative irrigation waters, i.e., Llobregat river water (LLRIV) and treated wastewater (SE: secondary effluent; WFPE: subsurface water flow-path effluent), included in this study.

	Irrigation Water	Lettuce	Strawberry	Parsley
Production Site 1	GW1.1 to GW1.4	LET1.1 to LET1.4	STR1.1 to STR1.8	PAR1.1 to PAR1.4
Production Site 2	GW2.1 to GW2.4	LET2.1 to LET2.4		
Production Site 3	RIV3.1 to RIV3.4	LET3.1 to LET3.4	STR3.1 to STR3.4	PAR3.1 to PAR3.4
Llobregat river	LLRIV.1 and LLRIV.2			
Treated wastewater	SE and WFPE			

Barcelona, with nearly 3.5 million inhabitants, has an important agricultural area on the outskirts of the city that uses water from the Llobregat river for irrigation. This river, which flows at 19 m³/s over a densely populated area (4948 km²), receives the effluents of more than 50 wastewater facilities and is affected by farming and agriculture activities. Agricultural fields at the Llobregat river delta cover an area of 3489 ha, and organic practices that use animal manure are increasing every year. Llobregat river water samples of 50 L ($n = 2$) were collected in autumn before the irrigation water intake (Table 1). Two treated wastewater samples (10 L), collected in winter, were also analyzed in this study, as this type of water is becoming an important source for irrigation, and it is also one of the main inputs of water from small river basins in the studied area (Table 1). The selected wastewater treatment plant (WWTP) is a conventional plant serving 147,000 inhabitants; it applies activated sludge as secondary treatment and a tertiary treatment based on *Phragmites australis*, a common species of reed that is autochthonous to the region and typically used in constructed wetlands and stream restoration actions [56], enabling the treatment of the secondary effluent in subsurface water flow paths receiving treated wastewater [57]. Secondary effluent ($n = 1$) and subsurface water flow-path effluent ($n = 1$) samples were collected, transported at 4 °C and processed on the same day of collection.

2.2. Viral Concentration and Nucleic Acid Extraction

The leafy sections of lettuce and parsley were cut into pieces with a length of approximately 2 cm and grouped into 25 g samples according to ISO 15216-1:2017 [38]. Each sample was washed for 20 min at 60 rpm in a rocker platform in Whirl-Pak® plastic bags

(Nasco, Fort Atkinson, WI, USA) containing 40 mL of Tris–Glycine–Beef Extract Buffer (pH 9.5, 0.25 N), and 0.25 volumes of 5× polyethylene glycol/NaCl solution was added. After precipitation for 1 h at 4 °C and centrifugation (10,000× g, 30 min, 4 °C), the resulting pellet was resuspended in 500 µL of PBS. Viral concentrates were treated with Turbo DNase (Invitrogen, Carlsbad, CA, USA) for 1 h at 37 °C to remove free DNA prior to nucleic acid extraction, and 500 µL of the DNase-treated viral concentrate was extracted using the NucliSENS® easyMAG system (BioMérieux, Marcy l’Etoile, France). Nucleic acids were eluted in 100 µL and stored at −80 °C for further analysis in qPCR and NGS assays.

In accordance with ISO 15216-1:2017 [38], 25 g of strawberries were treated with tris–glycine–beef extract buffer and pectinase at a pH of 9.5. The solution was centrifuged (8000× g, 10 min, 4 °C), and the pH was adjusted to 7.0 (±0.2) by using 0.1 N HCl. Then, 0.25 volumes of 5× polyethylene glycol/NaCl solution were added. After precipitation for 1 h at 4 °C and centrifugation (10,000× g, 30 min, 4 °C), the resulting pellet was resuspended in 500 µL of PBS and washed with chloroform/butanol. Nucleic acids were extracted from the resulting supernatant using the NucliSENS® easyMAG system (BioMérieux, Marcy l’Etoile, France) and stored at −80 °C for further analysis in qPCR and NGS assays.

Viruses were concentrated from irrigation water samples by ultrafiltration using the Long Volume Concentration kit (LVC kit) from InnovaPrep® (InnovaPrep, Drexel, MO, USA) that couples to a Rexeed 25-A polysulfone hollow-fiber ultrafilter. The elution step was performed using wet foam elution cans with Tris–PBS from InnovaPrep® (InnovaPrep, Drexel, MO, USA) and the eluted volume (30–50 mL) was further concentrated using Amicon Ultra-15 devices (50 kDa MWCO) from Millipore (Millipore, Burlington, MA, USA). Viral particles were recovered in a final volume of 200–500 µL of eluate and stored at −80 °C until further use. Viral concentrates were treated with Turbo DNase (Invitrogen, Carlsbad, CA, USA) for 1 h at 37 °C to remove free DNA prior to nucleic acid extraction, and 280 µL of the DNase-treated viral concentrate was extracted using the QIAamp® Viral RNA Mini Kit from QIAGEN (QIAGEN, Germantown, MD, USA). Nucleic acids were eluted in 80 µL and stored at −80 °C for further analysis in qPCR and NGS assays.

2.3. Viral Quantification

To evaluate the level of human fecal contamination in fresh food and irrigation water samples, a specific quantitative polymerase chain reaction (qPCR) assay for HAdV was performed using TaqMan® Environmental Master Mix 2.0 (Applied Biosystems, Waltham, MA, USA) with the specific primers and probe previously described [58]. Viral RNA pathogens were also quantified in all samples using the RNA UltraSense™ One-Step qRT-PCR System (Applied Biosystems, Waltham, MA, USA) with specific primers and probes for NoV GI [59–61], NoV GII [62,63], and HEV [64]. All assays were carried out using the Stratagene MX3000P sequence detector system (Agilent Technologies, Santa Clara, CA, USA). The qPCR standards were prepared using synthetic gBlocks® Gene fragments (IDT, Coralville, IA, USA), and serial dilutions were quantified using Qubit 3.0 dsDNA HS Assay Kit (Invitrogen, Carlsbad, CA, USA). All qPCR assays were performed in quadruplicate and included non-template controls.

LOD Determination

The limit of detection (LoD) of the qPCR was calculated by running six 10-fold dilutions of target DNA/RNA suspensions around the detection end points (2.5, 5, 25, and 50 Genome Copies (GC)/reaction) for each analyzed virus. The concentration that produced at least 95% positive replicates was assumed to be the LoD of the qPCR assay, which was transformed to the LoD of the entire process using the sample volume or grams tested in each analysis performed.

2.4. Target Enrichment Sequencing (TES)

2.4.1. Sequence-Independent, Single-Primer Amplification (SISPA)

A selection of samples that showed the presence of fecal contamination or specific viral pathogens were further analyzed using the TES approach, together with a negative control. Sample preparation prior to library construction consisted of the random tagging of nucleic acids and pre-amplification, allowing the study of both RNA and DNA viruses. This approach was performed following the procedure described previously [21,65,66], with the difference that SuperScript IV enzyme (Invitrogen, Carlsbad, CA, USA) was used in the retrotranscription step. Briefly, retrotranscription was performed using a random nonamer primer and followed by second-strand synthesis using Sequenase 2.0 (Applied Biosystems, Waltham, MA, USA). To obtain enough dsDNA for library construction, nucleic acids were amplified by 25 PCR cycles using AmpliTaqGold (Applied Biosystems, Waltham, MA, USA). The obtained PCR products were cleaned and concentrated with Zymo DNA Clean & Concentrate kit (Zymo Research, Irvine, CA, USA) and quantified using the Qubit 3.0 dsDNA HS Assay Kit (Invitrogen, Carlsbad, CA, USA).

2.4.2. Library Construction

For each sample, libraries were constructed using the KAPA HyperPlus Library Preparation Kit (KAPA Biosystems, Roche, Basel, Switzerland). Briefly, library construction consisted of fragmentation, indexation with KAPA Dual-Indexed Adapters (KAPA Biosystems, Roche, Basel, Switzerland) and amplification of the dsDNA obtained from SISPA. Afterwards, the resulting libraries were quantified using the Qubit 3.0 dsDNA HS Assay Kit (Invitrogen, Carlsbad, CA, USA).

2.4.3. Capture of Viral Sequences by VirCapSeq-VERT Capture Panel

Libraries were equimolarly pooled and captured using the VirCapSeqVERT Capture Panel (Roche, Basel, Switzerland). This panel involves the hybridization of probes designed to capture sequences from vertebrate viral pathogens, and it has enabled the detection of viral sequences in complex sample types in previous studies [21,49,50,67,68]. After the capture, quality and concentration were checked, and captured libraries were sequenced using an Illumina MiSeq 2 × 300 bp platform.

2.4.4. TES Bioinformatic Processing

Paired-end FASTAQ files generated from sequencing were analyzed using Genome Detective Virus Tool Version 1.126, a web-based software used to identify, assemble, and classify all known viruses present in NGS data (<https://www.genomedetective.com/app/typingtool/virus/>, accessed on 4 May 2021) [69]. For more precise and accurate taxonomic classification, human viral contigs obtained with nucleotide identity above 70% were further processed and queried for sequence similarity using BLASTN against the NCBI GenBank nucleotide collection database [70,71] with Geneious R9.1.8 (<https://www.geneious.com/>, accessed on 4 May 2021) [72].

NCBI Taxonomy standards were followed for the species nomenclature and classification. For specific typing of human caliciviruses, the obtained NoV contigs were further analyzed using the Noronet web-based Typing Tool (version 2.0) developed by RIVM using the updated classification of NoV genogroups and genotypes [73–75].

2.5. Amplicon Deep Sequencing (ADS)

2.5.1. Amplicon Generation

A selection of samples that showed the presence of fecal contamination or specific viral pathogens were further analyzed using an ADS approach. Sample nucleic acid suspensions were processed by ADS using specific nested PCR for NoV and HPV, previously described as suitable for typing purposes [21,52,76–79], with the incorporation of Illumina adapters in the nested primers. The obtained amplicons were purified from agarose gel using QIAquick

Gel Extraction (QIAGEN, Germantown, MD, USA) and sequenced with an Illumina MiSeq 2 × 300 bp platform.

2.5.2. ADS Bioinformatic Processing

Sequences generated from ADS were classified using BLASTN [71] (identity > 90%; coverage > 80%; length alignment: NoV > 200 bp, HPV > 100 bp) against a custom database populated with NoV prototype strain sequences from GenBank for the 9 genotypes of genogroup GI and 26 genotypes of genogroup GII determined by VP1, following the NoV classification proposed by Chhabra et al., 2019 [73] (Corrigendum 2020 [74]). The HPV database included sequences obtained from the International HPV Reference Center of Karolinska Institutet (<https://www.hpvcenter.se/>, accessed on 4 May 2021), consisting of all currently described HPV types, from HPV-1 to HPV-227. Other vertebrate papillomavirus genomes obtained from the Papillomavirus Episteme (PaVE) (<https://pave.niaid.nih.gov/>, accessed on 4 May 2021) were also included in the database.

3. Results

3.1. Virus Quantification in Irrigation Water and Organic Food

Samples of fresh fruits and vegetables from three organic farmers and irrigation water were tested for the presence of viral contaminants by specific qPCR assays. Low levels of contamination with human pathogenic viruses were detected in 46.9% of fresh produce samples and 50% of irrigation water samples (Table S1). The principal agent detected was NoV GII. Llobregat river water and treated wastewater samples showed simultaneous contamination with HAdV, NoV GI, and NoV GII, which were present in higher concentrations than in the rest of the samples.

The presence of HAdV, an indicator of human fecal contamination, was detected in 100% of the river water samples from the Llobregat river (LLRIV.1 and LLRIV.2) and 100% of the treated wastewater samples (SE and WFPE). Lower percentages of human fecal contamination were detected in the organic food samples: 37.5% of parsley samples, 33.3% of lettuce samples and 16.6% of strawberry samples analyzed from all production sites. No human fecal contamination was detected in groundwater samples using HAdV quantification. Figure 1 summarizes the mean viral concentrations measured in irrigation waters and harvested food samples from organic production sites by specific qPCR, and Table 2 specifies viral concentrations in alternative irrigation water sources. Further details can be found in Supplementary Materials, Table S1.

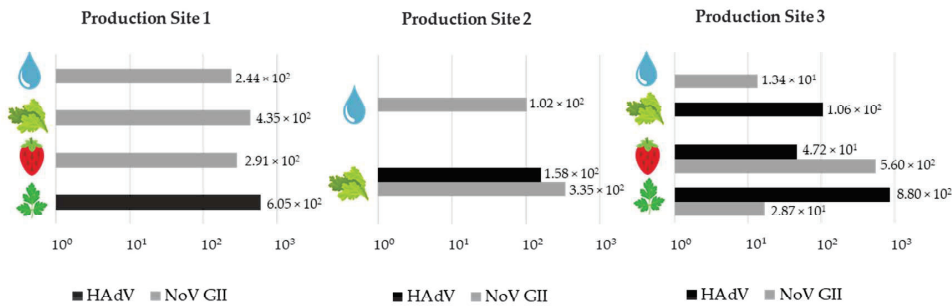


Figure 1. Quantification of viral pathogens in irrigation waters and harvested food samples (lettuce, strawberry, and parsley) from organic Production Sites 1, 2, and 3. Values on the X-axis show the mean concentrations of HAdV and NoV GII detected by qPCR assays, expressed in GC/25 g for food samples and in GC/L for irrigation water samples.

Table 2. Viral concentrations obtained from alternative irrigation water sources. Values are expressed in GC/L. ND: not detected. SE: secondary effluent, WFPE: subsurface water flow-path effluent.

		HAdV	NoV GI	NoV GII	HEV
Llobregat river	LLRIV.1	3.47×10^4	6.36×10^3	1.33×10^4	ND
	LLRIV.2	5.99×10^3	1.01×10^4	4.75×10^4	ND
Treated wastewater	SE	8.99×10^4	3.77×10^3	ND	ND
	WFPE	1.88×10^5	9.49×10^3	6.64×10^2	ND

NoV GI was found in 100% of Llobregat river water samples and 100% of treated wastewater samples, but NoV GI contamination was not detected in any of the irrigation water or food samples from the organic production sites. A different distribution was observed for NoV GII, which was widely present in all types of samples. This human viral pathogen was found in all irrigation water sources, including 50% of groundwater and river water samples, with the highest concentration detected in the Llobregat river sample. NoV GII was also detected in 25% of the organic food samples, with especially relevant values in all parsley samples collected from Production Site 3. HEV was not detected in any of the samples analyzed in this study.

The detection limits of the assays analyzing food samples were 75 GC/25 g for HAdV, 144 GC/25 g for NoV GI, 1036 GC/25 g for NoV GII, and 1000 GC/25 g for HEV. For the irrigation water samples, the detection limits of the assays analyzing groundwater and river water samples were 2.14 GC/L for HAdV, 4.11 GC/L for NoV GI, 29.60 GC/L for NoV GII, and 28.57 GC/L for HEV. In the analysis of treated wastewater samples, the limits of detection were 10.71 GC/L for HAdV, 20.57 GC/L for NoV GI, 148 GC/L for NoV GII, and 142.86 GC/L for HEV.

3.2. Virome of Irrigation Water and Organic Food Using TES

A selection of samples that showed the presence of fecal contamination or specific viral pathogens were further analyzed using the TES approach. These were: GW1.1, STR1.1, GW2.1, RIV3.1, LET3.4, PAR3.2, LLRIV.1, SE, and WFPE. This approach consisted of capturing sequences from vertebrate viral pathogens during library preparation using the VirCapSeqVERT Capture Panel (Roche, Basel, Switzerland). TES enabled the detection of human and other vertebrate viruses in all irrigation water samples and the identification of a human virus belonging to *Caliciviridae* in a parsley sample.

All contigs obtained from Genome Detective with nucleotide identity above 70% were further analyzed, resulting in a total of 77,565 viral reads: 31,634 reads for GW1.1, 44 reads for STR1.1, 977 reads for GW2.1, 1609 reads for RIV3.1, 20 reads for LET3.4, 9 reads for PAR3.2, 7446 reads for LLRIV.1, 16,421 reads for SE and 19,405 reads for WFPE. From the negative control included in the analysis, seven reads were assigned to the bacteriophage family *Microviridae*, but no other viral assignments were obtained. Further information about the assignment and distribution of all viral reads obtained in different hosts is detailed in Supplementary Materials, Table S2. Most of the reads belonged to bacteriophages (48.91%), followed by viruses that infect invertebrates (33.95%), plants (8.28%), humans (6.15%), and other vertebrates (2.44%). Using this capture approach, reads assigned to vertebrate viruses were obtained from all irrigation water samples, with the highest diversity found in a Llobregat river water sample (LLRIV.1). The distribution of vertebrate virus reads obtained using TES from different types of irrigation water sources is shown in Figure 2. Of the obtained reads of viruses that infect vertebrates, human viruses were present in greater proportions and accounted for all viruses detected in secondary-treated wastewater. In the Llobregat river, the highest number and diversity of reads were assigned to viruses that infect non-human vertebrates, including feline, canine, porcine, rodent, and cattle viruses belonging to *Astroviridae*, *Caliciviridae*, *Parvoviridae*, and *Picornaviridae*. Avian, equine and bat viruses belonging to *Genomoviridae* and *Parvoviridae* were detected in

groundwater, whereas river water from Production Site 3 was found to have feline and porcine viral members of *Astroviridae* and *Genomoviridae*. Treated wastewater showed a huge proportion of human virus reads, with a low proportion representing canine and cattle parvoviruses in the effluent of subsurface water flow paths.

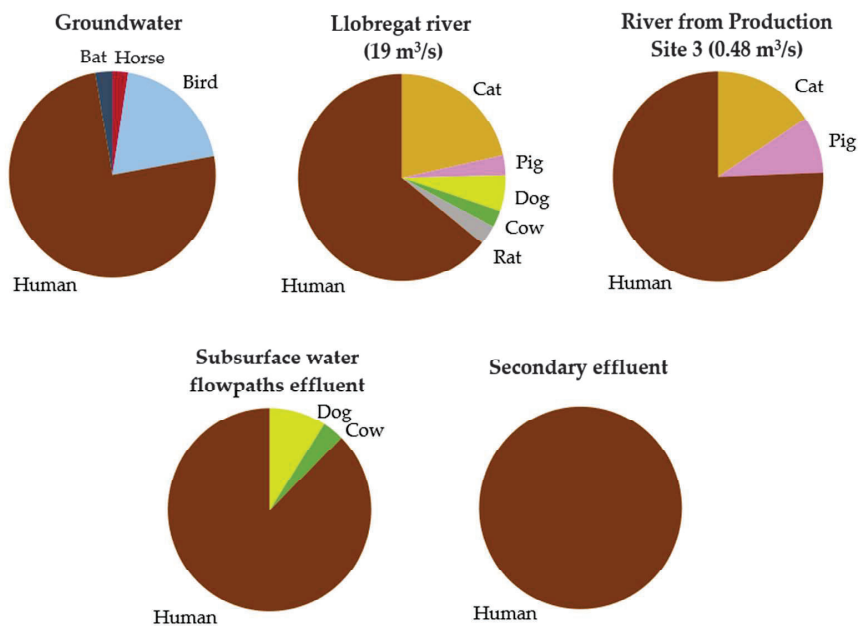


Figure 2. Distribution of vertebrate virus reads among different hosts obtained using target enrichment sequencing from different types of irrigation water sources: groundwater, Llobregat river (flow of 19 m³/s), river water from Production Site 3 (flow of 0.48 m³/s), subsurface water flow-path effluent and secondary effluent. Further details are provided in Supplementary Materials, Table S2.

In food samples, no reads of viruses that infect humans or other vertebrates were obtained from strawberry and lettuce samples, despite the detection of NoV GII and HAdV by qPCR, respectively, and three reads assigned to a human virus belonging to *Caliciviridae* were obtained from the parsley sample, which showed both HAdV and NoV GII contamination in qPCR assays.

The human pathogenic viruses sequenced using TES are summarized in Table 3. Despite having samples positive for HAdV in the qPCR assay, no assignments belonging to *Adenoviridae* were obtained. Reads assigned to members of *Caliciviridae* were obtained from the Llobregat river and the parsley sample, as mentioned previously. Sequences obtained from the Llobregat river were assigned to NoV GI.1, GII.4, and GII.17, the last of which was also the genotype found in parsley. In this study, all sequences assigned to *Astroviridae* were mainly detected in river water samples. The Llobregat river contained a wide diversity of Human Astroviruses (HAstVs) belonging to different species, including *Mamastrovirus 1* (HAstV-1 and HAstV-4), *Mamastrovirus 6* (AstV-MLB1 and AstV-MLB2), and *Mamastrovirus 8* (HAstV-VA2). HAstV-1 and HAstV-5 were also detected in the river sample from Production Site 3. The presence of HAstV-5 was also detected in subsurface water flow-path effluent. Members of *Picornaviridae* were detected in river water and treated wastewater. Sequences of *Aichivirus A*, typed as Aichi virus 1, were sequenced from river water, whereas *Salivirus A* was found in river water and treated wastewater. Other assignments obtained from irrigation water samples belonged to *Circoviridae*, with sequences detected in groundwater and secondary effluent, and *Parvoviridae*, with sequences detected in groundwater, river water, and subsurface water flow-path effluent.

Table 3. Human viruses sequenced in irrigation water from Production Site 1, Production Site 2, and Production Site 3, alternative irrigation water sources and parsley from Production Site 3 analyzed using target enrichment sequencing.

Family	Genus	Species	Genogroup/ Genotype/ Serotype	Samples					
				Contigs	Length (bp)	Nucleotide Identity (%)	Genome Coverage (%)	Sample Name	Site
Astroviridae	Mamastrovirus	Mamastrovirus 1	HAstV-1	2	429–435	98.60–99.07	6.3–6.4	RIV3.1	Production Site 3
				1	551	98.19	8.1	LLRIV.1	Llobregat river
				2	800–3071	92.10–93.62	11.9–45.7	LLRIV.1	Llobregat river
		Mamastrovirus 6	HAstV-5	1	2869	92.36	42.2	RIV3.1	Production Site 3
				1	808	97.65	12.1	WFPE	Treated wastewater
				1	420	95.00	6.8	LLRIV.1	Llobregat river
Mamastrovirus 8	HAstV-VA2	1	364	98.35	5.9	LLRIV.1	Llobregat river		
		1	585	99.49	9.0	LLRIV.1	Llobregat river		
		1	524	90.15	6.8	LLRIV.1	Llobregat river		
Caliciviridae	Norovirus	Norwalk virus	NoV GI.4	2	689–845	92.79–97.98	9.1–11.2	LLRIV.1	Llobregat river
				1	282	97.28	3.8	PAR3.2	Production Site 3
				2	632–717	95.66–98.58	8.4–9.5	LLRIV.1	Llobregat river
Circoviridae	Cyclovirus	Human associated cyclovirus 6	NG12	1	1541	81.00	85.9	GW1.1	Production Site 1
				1	433	88.29	24.10	SE	Treated wastewater
				1	124	95.97	2.7	GW2.1	Production Site 2
Parvoviridae	Dependoparvovirus	Adeno-associated dependoparvovirus A	AAV2	1	3040	97.27	64.9	RIV3.1	Production Site 3
				3	932–1216	96.46–98.39	19.9–26.0	LLRIV.1	Llobregat river
				1	1438	96.45	30.7	WFPE	Treated wastewater
		Bocaparvovirus	AAV6	1	682	79.71	14.6	LLRIV.1	Llobregat river
				3	342–566	95.61–98.25	6.6–10.8	LLRIV.1	Llobregat river
				4	348–724	94.31–99.42	4.2–8.8	RIV3.1	Production Site 3
Picornaviridae	Kobuvirus	Aichi virus A	Aichi virus 1	5	404–952	95.79–97.56	4.9–11.5	LLRIV.1	Llobregat river
				3	683–1197	97.21–98.41	8.7–15.3	RIV3.1	Production Site 3
				2	459–552	97.67–98.04	5.9–7.0	LLRIV.1	Llobregat river
	Salivirus	Salivirus A	Undetermined	2	395–907	96.73–97.97	5–11.6	SE	Treated wastewater
				3	516–623	97.16–97.87	6.6–7.9	WFPE	Treated wastewater

NoV sequences obtained using ADS and classified by VP1 analysis are shown in Table 4. Treated wastewater samples (SE and WFPE) presented the highest diversity of NoV GI, whereas the highest diversity of NoV GII was detected in Llobregat river water (LLRIV.1). NoV GI.4 was the most abundant genotype, and it also accounted for the highest number of NoV reads in Llobregat river water. Among NoV GII sequences, NoV GII.4 and GII.13 were the genotypes that appeared in a greater number of samples, and a high number of reads were obtained mainly from Llobregat river water. Other NoV GII genotypes sequenced from the Llobregat river were GII.2, GII.17, and GII.3.

Table 4. Number of reads of each NoV genotype obtained from a strawberry sample from Production Site 1, irrigation water and parsley samples from Production Site 3 and alternative irrigation water sources using amplicon deep sequencing.

		Production Site 1	Production Site 3			Llobregat River		Treated Wastewater		
		STR1.1	RIV3.1	PAR3.1	PAR3.2	LLRIV.1	LLRIV.2	SE	WFPE	
NoV GI	GI.4	17	1		10	98,699	3	15	59	
	GI.1	4	2			15		23,411	35,901	
	GI.2					8	28	2	2	
	GI.3					4		9	9500	
	GI.5		2					12,021	475	
NoV GII	GII.4		1		3	9331		3	6	
	GII.13		2	3	11	49,651			16	
	GII.2			2		16,933			4	
	GII.17					52	1203			
	GII.3					68				
			1-10	10-100	100-1000	1000-10,000	10,000-100,000	>100,000		

All food samples presented a lower diversity of NoV, but it is remarkable that all NoV genotypes identified were also found in some of the irrigation water sources. NoV GI members found in food samples, GI.4 and GI.1 in strawberry (STR1.1), and GI.4 in parsley (PAR3.2), were also found in river water and treated wastewater. NoV GII genotypes found in parsley samples, GII.13 and GII.2 in PAR3.1 and GII.4 and GII.13 in PAR3.2, were also found in different irrigation water sources.

HPV and other PV sequences obtained using ADS and classified by L1 analysis are shown in Table 5. The sequences obtained were classified into 16 different HPV types, most of which were members of the genus *Betapapillomavirus*. Llobregat river water and treated wastewater showed the highest diversity of HPV. HPV-92, HPV-105, and HPV-122, belonging to the species *Betapapillomavirus* 4, 1, and 2, respectively, were sequenced from all samples. Few *Alphapapillomavirus* were identified (HPV-177 and HPV-57), and only one *Gammapapillomavirus* (HPV-4) was found in river and lettuce samples from Production Site 3, in the Llobregat river and in treated wastewater samples (SE and WFPE). Other PVs not classified as HPV were sequenced: *Bos taurus* *Papillomavirus* 7 (BPV-7) from treated wastewater samples and *Rattus norvegicus* *Papillomavirus* 2 (RnPV-2) from a Llobregat river sample (LLRIV.2).

Table 5. Number of reads of each type of human and other Papillomavirus obtained from irrigation water, lettuce, and parsley samples from Production Site 3 and alternative irrigation water sources using amplicon deep sequencing.

		Production Site 3				Llobregat River		Treated Wastewater	
		RIV3.1	LET3.4	PAR3.1	PAR3.2	LLRIV.1	LLRIV.2	SE	WFPE
HPV	HPV-92	1	2	14	6	39,158	15	3	16
	HPV-105	3	45	381,617	16	108	197	13	337,323
	HPV-122	2	19	90	50	271,741	38	11	114
	HPV-38		152,893	14	19	23	1161	8	7
	HPV-145		102,028	17	8	8	150	5	6
	HPV-177	16	10	3		10	20	8	28
	HPV-182	1	9	580		1	1	3	2
	HPV-37		101		45	8	9	90	19
	HPV-76		84	6		8	87,512	4	10
	HPV-96		2			1	145	4	1171
	HPV-4	7	1				2	169,019	3
	HPV-57	6	56					76,118	1
	HPV-12		20	16		79	152,261	17	35
	HPV-8							321	1
	HPV-17						2		3
Other PV	HPV-196						49	1	
	BPV-7							1	1
	RnPV-2						17		
		1-10	10-100	100-1000	1000-10,000	10,000-100,000	>100,000		

As was the case with NoV, all HPV types detected in food samples were also identified in some of the irrigation waters.

4. Discussion

At present, quantitative PCR is the most common method applied to the identification of viral contamination in foods. ISO methods focused on qPCR detection and quantification of NoV and HAV in different food matrices are available and can be applied even when viral food quality is not included in any regulations. Additionally, HAdV has been suggested and proven to be useful as a viral fecal indicator in water matrices, as this human pathogen is widely detected when water is affected by sewage [80–83] and has been previously detected in conventional and reclaimed irrigation water [29]. In this study, qPCR detected the presence of HAdV in river water and treated wastewater, indicating that irrigation water is a potential source of human fecal contamination and can potentially become a vehicle for viruses that are transmitted via the fecal–oral route through fresh fruits and vegetables. Using the HAdV indicator, fecal contamination was also detected in the parsley, lettuce, and strawberry samples tested in this study. In 2017, Fernández-Cassi et al. also described the presence of this indicator in water samples from the Besòs river and parsley samples irrigated with this river water source [5]. Other studies have reported HAdV in lettuce and irrigation water [84] and in strawberries collected from European food production chains [85].

In addition to HAdV, other pathogenic viruses were detected by qPCR in this study. NoV GI was found in river water and treated wastewater, but it was not detected in food or irrigation water samples from the organic production sites, despite the fact that some were positive for HAdV. Other studies reported NoV GI contamination in strawberries and suspected irrigation water to be the potential source of contamination [86]. On the contrary, NoV GII was widely present in all types of samples. All irrigation water sources and 25%

of food samples showed NoV GII contamination in concentrations consistent with those reported in other studies in which polluted water was suspected to be the contamination source of NoV in fresh fruits and vegetables [84,85,87]. Secondary effluent and the effluent of subsurface water flow paths showed similar levels of contamination, indicating that these flow paths were not efficient in eliminating viral contaminants.

HEV was not detected in any of the samples analyzed in this study. Other studies have found HEV to be present in fresh produce: in a pack of frozen raspberries taken from a cold room at point-of-sale [85], in an irrigated, field-grown strawberry sample suspected to be contaminated by irrigation water [86] and in lettuce heads from primary production sites and at point-of-sale [84]. This last study also reported the presence of HEV in irrigation groundwater, which was also described by Rusiñol et al. in an area with intensive pig farming activities [22]. HEV is a zoonotic pathogen that can cause self-limiting or fulminant hepatitis in humans, and it is important to check for its presence since an increasing number of foodborne HEV cases are being reported in Europe, frequently associated with the consumption of pork products [27].

Ultimately, qPCR assays revealed low but frequent contamination associated with human pathogenic viruses in fresh produce and irrigation water, with NoV GII being the principal agent detected. Regarding the NGS data obtained in this study, a targeted assay (TES) was applied by using the VirCapSeqVERT Capture Panel (Roche, Basel, Switzerland) on a selection of samples showing fecal contamination or the presence of specific viral pathogens. This panel employs approximately 2 million biotinylated oligonucleotide probes designed to bind coding sequences of all viral taxa known to infect vertebrates. The TES approach is known to facilitate the detection of vertebrate viruses [68], although some sequences from other viral hosts can be identified [21]. In fact, with the real picture of the virome in mind, vertebrate viruses are commonly found in low proportions with respect to other viruses such as bacteriophages and plant viruses, which have been described as the most abundant in sewage samples [65,88]. TES has been successfully employed in previous studies [21,49,50,67,68] to improve the detection of vertebrate viruses of interest that would otherwise be difficult due to their low concentrations in environmental samples. Recently, the VirCapSeqVERT Capture Panel was applied in a metagenomic study to evaluate NoV genomic diversity in oysters [50], but to date, no studies have applied this capture approach to fresh fruits and vegetables. In this study, TES allowed the acquisition of reads assigned to vertebrate viruses from all irrigation water samples, with the highest diversity observed in the Llobregat river water sample (LLRIV.1). Among the viruses identified in food samples, NoV GII sequences obtained from parsley, which showed a concentration of 7.12×10^1 GC/25 g by NoV GII qPCR, constituted a unique contig that was typed as NoV GII.17. This genotype emerged as a major cause of gastroenteritis outbreaks in China and Japan in the winter of 2014/2015 [89]. Interestingly, NoV GII sequences were not obtained from the strawberry sample STR1.1 using the capture approach despite it showing a high concentration of NoV GII in the specific qPCR assay. This result could be associated with differences in food matrices: detecting viruses in berries is known to be challenging due to the presence of various inhibitory substances and low pH [90]. Bartsch et al. also obtained a low number of NoV reads from the metagenomic analysis of frozen strawberries involved in a large NoV gastroenteritis outbreak [39]. The emergent genotype detected in parsley—NoV GII.17—was also sequenced from the Llobregat river, which also showed other sequences belonging to the *Norwalk virus* typed as NoV GI.1 and NoV GII.4.

ADS assays enabled a better characterization of the NoV diversity present in the studied fresh produce and irrigation water samples. This approach, based on the mass sequencing of traditionally Sanger-sequenced PCR amplicons, facilitates the detailed study of specific families and their diversity within a sample [21] and has previously been applied to environmental and shellfish samples for studying viral groups—such as *Adenoviridae* [21,65,91], *Papillomaviridae* [21,23,92,93], or NoVs [48,52,53]—but none of these studies investigated fresh food samples. Due to differences in the sensitivities of NoV assays and the fact that the concentrations detected were low, not all qPCR-positive samples were

positive in RT-PCR amplification for ADS. In particular, STR1.1, PAR3.1, PAR3.2, and RIV3.1, which were qPCR-positive for NoV GII and qPCR-negative for NoV GI, produced an RT-PCR amplicon of the expected size for NoV GI but not for NoV GII. Additionally, groundwater and treated wastewater (WPFE) samples were positive for NoV GII in the qPCR assay, but no amplicons of the expected size were produced by conventional RT-PCR. This result may be explained by the lower sensitivity of this technique as compared to the qPCR assay [87,94]. In contrast, LET3.4 showed an amplicon of the expected size for NoV GI, although it did not show (or at least not above the limit of detection) NoV contamination by qPCR, which could be due to sensitivity differences between assays when concentrations are around this limit.

NoV GII.4 has been reported to be the most predominant genotype worldwide [95,96]. This is consistent with our findings, in which the presence of the GII.4 genotype was widely observed in different samples using ADS, including river water and treated wastewater samples, and also in a parsley sample. Llobregat river water was the sample with the highest number of sequences obtained from this specific genotype using ADS, and it was also the only sample in which this genotype was possible to sequence using TES. In fact, the three genotypes identified from the Llobregat river using TES, GI.1, GII.4, and GII.17, were also identified by ADS. It is important to highlight that all NoV genotypes found in food samples, GI.4 and GI.1 in strawberry and GI.4, GII.4, GII.13, and GII.2 in parsley, were also found in irrigation water samples. Maunula et al. also detected NoV GI.4 in frozen raspberries from one particular batch that was confirmed to be the source of a described cluster of NoV outbreaks affecting about 200 people in Southern Finland in 2009 [97].

In addition to viruses belonging to *Caliciviridae*, the TES results also showed other pathogenic or potentially pathogenic viruses present in river water, specifically viruses belonging to *Astroviridae*, *Picornaviridae*, and *Parvoviridae* families, which is consistent with the virome previously described in a river in the same geographical area that is also used for irrigation [5]. The Llobregat river contained a wide diversity of HAsTVs, including the species *Mamastrovirus* 1, 6, and 8. HAsTVs are important agents causing acute gastroenteritis in children and have been involved in outbreaks affecting adults [95,98]. Although HAsTV prevalence seems to be of lower importance compared to the number of outbreaks caused by NoV, its importance might be underestimated [5]. Recently, HAsTVs have also been associated with other pathologies, such as meningitis and acute flaccid paralysis [95,99–101]. HAsTV-1, which was found in river water samples, is the most common type associated with infantile gastroenteritis [95,102]. *Picornaviridae* sequences found in river water samples belonged to the genus *Kobuvirus*, specifically to *Aichivirus A*, a viral species that includes members recognized as human pathogens that cause gastroenteritis outbreaks [103–105]. Other *Picornaviridae* sequences were assigned to *Salivirus A*, belonging to the genus *Salivirus*. The highest proportion of the vertebrate virus sequences obtained were human viruses, and among all types of irrigation water, the Llobregat river showed the highest number of reads and diversity of viruses that infect non-human vertebrates, including feline, canine, porcine, rodent, and cattle viruses belonging to *Astroviridae*, *Caliciviridae*, *Parvoviridae*, and *Picornaviridae* families. River water from Production Site 3 presented feline and porcine viruses designated as *Astroviridae* members. Among all vertebrate virus sequences obtained using TES, the greatest proportion of viruses that infect humans was observed, as expected, in treated wastewater, with human viruses accounting for all viruses detected in secondary-treated wastewater, and were identified as members belonging to *Astroviridae*, *Circoviridae*, *Parvoviridae*, and *Picornaviridae*, which is consistent with other metagenomic studies [22,106]. The non-human vertebrate viruses sequenced in treated wastewater were canine and cattle parvoviruses. Human viral sequences belonging to *Circoviridae* and *Parvoviridae* and avian, equine and bat viruses belonging to *Genomoviridae* and *Parvoviridae* were also detected in groundwater samples. Viruses are commonly detected in groundwater used for irrigation, even though it is often considered a microbially safe source [7,107]. These viruses may originate from the leakage of sewage water or diffuse contamination from livestock production zones close to produce fields [108]. In this study, the sequences

obtained using TES assigned to viruses that infect non-human vertebrates of interest could be related to intensive farming activities in the surrounding areas, which have the potential to contaminate irrigation water through leakage. Despite lower yields of production in organic farming compared with intensive agriculture, organic farming systems represent an attractive environmentally friendly alternative, delivering equally or more nutritious foods that contain less, or no, pesticide residues [109]. This system relies on the incorporation of organic material into the soil using animal manure as fertilizer [110–112]. Although animal manure is considered a beneficial organic fertilizer and a good source of nutrients, it is also a well-known source of foodborne pathogenic bacteria, parasites, and viruses if it is not adequately aged or treated before application [13,113,114] and, therefore, is a potential consumption risk.

Despite having samples positive for HAdV in the qPCR assay, no assignments belonging to *Adenoviridae* were obtained using TES. It has been previously reported that the SISPA protocol, which is performed prior to library preparation in order to overcome the limitation of low quantities of viral genomes, might introduce bias by amplifying the most abundant genomes and, therefore, underrepresenting others [115,116]. Fernández-Cassi et al. already noted these specific difficulties when detecting low numbers of dsDNA viruses such as HAdV in river water or when unable to detect HAdV in parsley samples that were positive for the virus in the qPCR assay [5]. Considering that TES did not allow the acquisition of HAdV sequences, a more suitable approach, such as ADS, should be applied to improve the sensitivity of viral metagenomics. Similarly, no sequences of *Papillomaviridae* were obtained from the TES approach. The detection of HPVs could also be affected by SISPA bias due to their dsDNA genome. However, ADS was demonstrated to be a suitable HPV detection and diversity exploration tool that overcomes this specific TES limitation since a wide variety of HPV was observed using this approach. Two different primer sets were used: GP5+/GP6+ were designed for detecting mucosal HPV types [79], whereas FAP6085/6319 were designed for detecting cutaneous HPV types [78]. The role of cutaneous HPV types in pathogenesis remains unclear: some are regarded as potential high-risk types because they are found in squamous cell carcinoma (SCC), such as HPV-5 or HPV-8, while most other cutaneous HPVs are only associated with benign lesions such as epidermodysplasia verruciformis [117,118]. Most of the HPV types sequenced in this study were classified as *Betapapillomavirus*. River and treated wastewater samples contained the highest diversity of *Betapapillomavirus*: HPV-12, HPV-17, HPV-37, HPV-38, HPV-76, HPV-92, HPV-96, HPV-105, HPV-122, HPV-145, and HPV-182. Treated wastewater also contained HPV-196 and HPV-8 sequences, the last of which is the HPV mostly found in SCC of the skin [117], and it has been described as an abundant HPV type in urban sewage [21]. HPV-12, HPV-37, HPV-38, HPV-76, HPV-92, HPV-105, HPV-122, and HPV-145 types were also identified in lettuce and parsley samples, and HPV-96 was only detected in lettuce. HPVs belonging to *Alphapapillomavirus* were also sequenced. HPV-177 was identified in river water and treated wastewater, as well as in lettuce and parsley samples. HPV-57 is associated with common warts typically occurring in the anogenital region [119] and was identified in river water, treated wastewater, and lettuce. Members belonging to the genus *Alphapapillomavirus* have been described previously in raw sewage [21,23]. HPV was frequently detected in the food and irrigation water samples analyzed in this study and should therefore be considered a potential emergent pathogen, and the role of irrigation water and fresh produce in the transmission of HPV should be further investigated. To our knowledge, this study is the first to provide data on the diversity of HPVs and NoVs present in fresh fruit and vegetable samples. Manipulation during the production process should not be ruled out as a route through which viral contamination is introduced to the food analyzed in this study. It is possible that some of the HPVs detected in food samples were skin contaminants, since most of the reads were classified as members belonging to the genus *Betapapillomavirus*, which includes types commonly isolated from skin. From sample collection to analysis, samples were handled carefully using gloves during all processes in order to avoid any kind of viral contamination. It is important to remark that

all viruses detected in food were also detected in water samples, supporting the possibility that irrigation water is a source of contaminant HPVs.

The results of this study indicate that irrigation water and fresh fruits and vegetables present a wide variety of viral pathogens that may pose a risk to humans. Natural water treatments or disinfection procedures are often necessary to obtain high quality irrigation water. The implementation of quality monitoring programs integrated with Quantitative Microbial Risk Assessment (QMRA) investigations will provide an estimate of the level of risk and the treatments needed to produce high quality water. The NGS techniques described in this study would be useful for the identification of significant pathogens present in water and food for risk assessment studies, the selection of water treatments and the development of optimized site-specific safety plans. Because wastewater is usually of poor chemical and microbiological quality, extensive treatments are required before it can be safely used for irrigation [113]. However, drought conditions are linked to the utilization of untreated wastewater or contaminated groundwater for irrigation due to the lack of clean water, increasing the chances of microbial contamination of food products and soil [120]. In addition, due to climate change, unpredictable meteorological phenomena are expected to become more frequent and likely to drive an increase in the incidence of water scarcity with a probability of drought [1], potentially leading to the utilization of reclaimed water as a common source for irrigation in many locations. The association between pathogen diversity in this type of irrigation water and foodborne disease would imply that efficient treatments must be implemented to promote its use as an additional and safe water source. This study only suggests a possible link between irrigation water and the contamination origin of fresh produce, but other contamination factors involved in agricultural practices could certainly have an impact and should be considered, such as organic fertilizer or run-off from nearby animal pastures, which could also be vehicles for viral contamination [85,121]. Furthermore, detection of viral genomic sequences does not necessarily imply a consumption risk since infectivity potential was not verified in this study. In sum, irrigation water and animal manure are considered the two most important pathways of pathogen transmission from human or animal hosts to fresh produce at the preharvest level [10].

Checking for the presence of viruses as a part of controlling the quality of irrigation water is a key intervention step to reduce the risk of transferring contamination to fresh produce [108,122]. Further studies are needed to confirm and elucidate the significance of the information provided in terms of risk for consumption. From the data obtained in this study, we conclude that TES is a useful tool to obtain a broad picture of vertebrate viruses that integrate into the virome of irrigation water samples, whereas ADS allows the in-depth characterization of the diversity of a specific viral pathogen contaminating irrigation water and fresh produce. More relevant is that, by applying these methodologies, human pathogenic viruses were detected in samples in which qPCR showed low levels of contamination, suggesting that NGS approaches could be a suitable tool to identify and characterize viral pathogens and improve food monitoring and foodborne disease outbreak control.

5. Conclusions

- The fresh fruits and vegetables cultured in the analyzed organic farms show a high frequency of viral contamination, and the contaminating viruses are also detected in river water and tertiary effluents from subsurface water flow paths used in irrigation, indicating that they are common circulating pathogens.
- In this study, irrigation water is the most probable source of the viral pathogens, primarily NoV and HPV, detected in food. All pathogens were detected in some types of irrigation water, such as river water or treated wastewater, indicating that are potential sources of contamination. However, virus infectivity potential was not analyzed in this study.

- A wide diversity of cutaneous HPV was detected in fresh produce and a wider diversity of HPV was identified in more polluted water samples, such as river and treated wastewater samples. HPV-8, a high-risk type associated with SCC, was found in treated wastewater.
- The three methodologies used for the analysis of viral contamination of irrigation water and organic food are useful and produce different types of information: (a) qPCR is a highly sensitive quantitative and specific technique, (b) TES shows the presence of relatively abundant viral pathogens present in irrigation water, including unexpected viral pathogens and potentially zoonotic strains, and (c) ADS provides higher sensitivity for the identification of viral types or variants in viral groups that contaminate food and water.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/foods10081820/s1>, Table S1: Quantification of viral pathogens in irrigation water and fresh produce samples. Values are expressed in GC/L for irrigation water samples and in GC/25 g for food samples. ND: Not Detected; Table S2: Reads obtained from each viral assignment in irrigation water, strawberry, lettuce, and parsley samples using target enrichment sequencing. Colors represent the specific host of vertebrate viruses, legend can be found under the table.

Author Contributions: Conceptualization, R.G., M.R., S.B.-M., S.M.-P. and M.I.; Methodology, M.I., R.G., S.M.-P., E.F., M.R. and A.H.; Software, S.M.-P., M.I. and N.T.; Validation, M.R., S.B.-M. and R.G.; Formal analysis, M.R., R.G., M.I., S.M.-P. and N.T.; Investigation, M.I. and M.R.; Resources, S.B.-M. and R.G.; Data curation, M.I., M.R., R.G. and S.B.-M.; Writing—original draft preparation, M.I., M.R. and R.G.; Writing—review and editing, M.I., M.R., S.B.-M., S.M.-P., E.F., N.T., A.H. and R.G.; Visualization, M.I., M.R., S.M.-P. and S.B.-M.; Supervision, R.G. and M.R.; Project administration, M.R., S.B.-M. and R.G.; Funding acquisition, M.R., S.B.-M. and R.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially funded by a grant from the Spanish Ministry of Economy and Competitiveness (MINECO) AGL2017-86797-C2-1-R. IDEA-CSIC is a Centre of Excellence Severo Ochoa (Spanish Ministry of Science and Innovation (MICINN), Project CEX2018-000794-S).

Data Availability Statement: The datasets generated during the current study are available in Zenodo under the DOI number <https://doi.org/10.5281/zenodo.4722797>.

Acknowledgments: Sílvia Bofill-Mas is a Serra-Hunter fellow at the University of Barcelona. Eva Forés is an APIF fellow at the University of Barcelona. Marta Itarte is a fellow of the Catalan Government “AGAUR” (FI) at the University of Barcelona.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hess, T.; Sutcliffe, C. The Exposure of a Fresh Fruit and Vegetable Supply Chain to Global Water-Related Risks. *Water Int.* **2018**, *43*, 746–761. [\[CrossRef\]](#)
2. Balali, G.I.; Yar, D.D.; Afua Dela, V.G.; Adjei-Kusi, P. Microbial Contamination, an Increasing Threat to the Consumption of Fresh Fruits and Vegetables in Today’s World. *Int. J. Microbiol.* **2020**, *2020*, 3029295. [\[CrossRef\]](#)
3. Septembre-Malaterre, A.; Remize, F.; Poucheret, P. Fruits and Vegetables, as a Source of Nutritional Compounds and Phytochemicals: Changes in Bioactive Compounds during Lactic Fermentation. *Food Res. Int.* **2018**, *104*, 86–99. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Mir, S.A.; Shah, M.A.; Mir, M.M.; Dar, B.N.; Greiner, R.; Roohinejad, S. Microbiological Contamination of Ready-to-Eat Vegetable Salads in Developing Countries and Potential Solutions in the Supply Chain to Control Microbial Pathogens. *Food Control* **2018**, *85*, 235–244. [\[CrossRef\]](#)
5. Fernandez-Cassi, X.; Timoneda, N.; Gonzales-Gustavson, E.; Abril, J.F.; Bofill-Mas, S.; Girones, R. A Metagenomic Assessment of Viral Contamination on Fresh Parsley Plants Irrigated with Fecally Tainted River Water. *Int. J. Food Microbiol.* **2017**, *257*, 80–90. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Callejón, R.M.; Rodríguez-Naranjo, M.I.; Ubeda, C.; Hornedo-Ortega, R.; Garcia-Parrilla, M.C.; Troncoso, A.M. Reported Foodborne Outbreaks Due to Fresh Produce in the United States and European Union: Trends and Causes. *Foodborne Pathog. Dis.* **2015**, *12*, 32–38. [\[CrossRef\]](#)
7. Shaheen, M.N.F.; Elmahdy, E.M.; Chawla-Sarkar, M. Quantitative PCR-Based Identification of Enteric Viruses Contaminating Fresh Produce and Surface Water Used for Irrigation in Egypt. *Environ. Sci. Pollut. Res.* **2019**, *26*, 21619–21628. [\[CrossRef\]](#)

8. Iwu, C.D.; Okoh, A.I. Preharvest Transmission Routes of Fresh Produce Associated Bacterial Pathogens with Outbreak Potentials: A Review. *Int. J. Environ. Res. Public Health* **2019**, *16*, 4407. [\[CrossRef\]](#)
9. Garcia, B.C.B.; Dimasupil, M.A.A.Z.; Vital, P.G.; Widmer, K.W.; Rivera, W.L. Fecal Contamination in Irrigation Water and Microbial Quality of Vegetable Primary Production in Urban Farms of Metro Manila, Philippines. *J. Environ. Sci. Health* **2015**, *50*, 734–743. [\[CrossRef\]](#)
10. Park, S.; Szonyi, B.; Gautam, R.; Nightingale, K.; Anciso, J.; Ivanek, R. Risk Factors for Microbial Contamination in Fruits and Vegetables at the Preharvest Level: A Systematic Review. *J. Food Prot.* **2012**, *75*, 2055–2081. [\[CrossRef\]](#)
11. Wei, X.; Hou, S.; Pan, X.; Xu, C.; Li, J.; Yu, H.; Chase, J.; Atwill, E.R.; Li, X.; Chen, K.; et al. Microbiological Contamination of Strawberries from U-Pick Farms in Guangzhou, China. *Int. J. Environ. Res. Public Health* **2019**, *16*, 4910. [\[CrossRef\]](#)
12. Rodrigues, C.; da Silva, A.L.B.R.; Dunn, L.L. Factors Impacting the Prevalence of Foodborne Pathogens in Agricultural Water Sources in the Southeastern United States. *Water* **2020**, *12*, 51. [\[CrossRef\]](#)
13. Gil, M.I.; Selma, M.V.; Suslow, T.; Jacksens, L.; Uyttendaele, M.; Allende, A. Pre- and Postharvest Preventive Measures and Intervention Strategies to Control Microbial Food Safety Hazards of Fresh Leafy Vegetables. *Crit. Rev. Food Sci. Nutr.* **2015**, *55*, 453–468. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Pachepsky, Y.; Shelton, D.R.; McLain, J.E.T.; Patel, J.; Mandrell, R.E. Irrigation Waters as a Source of Pathogenic Microorganisms in Produce. A Review. *Adv. Agron.* **2011**, *113*, 73–138. [\[CrossRef\]](#)
15. Machado-Moreira, B.; Richards, K.; Brennan, F.; Abram, F.; Burgess, C.M. Microbial Contamination of Fresh Produce: What, Where, and How? *Compr. Rev. Food Sci. Food Saf.* **2019**, *18*, 1727–1750. [\[CrossRef\]](#)
16. Gobeil, A.; Maherani, B.; Lacroix, M. Norovirus Elimination on the Surface of Fresh Foods. *Crit. Rev. Food Sci. Nutr.* **2020**. [\[CrossRef\]](#)
17. Gonzales-Gustavson, E.; Rusiñol, M.; Medema, G.; Calvo, M.; Girones, R. Quantitative Risk Assessment of Norovirus and Adenovirus for the Use of Reclaimed Water to Irrigate Lettuce in Catalonia. *Water Res.* **2019**, *153*, 91–99. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Dicaprio, E.; Purgianto, A.; Li, J. Effects of Abiotic and Biotic Stresses on the Internalization and Dissemination of Human Norovirus Surrogates in Growing Romaine. *Appl. Environ. Microbiol.* **2015**, *81*, 4791–4800. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Koo, H.L.; Ajami, N.; Atmar, R.L.; DuPont, H.L. Noroviruses: The Leading Cause of Gastroenteritis Worldwide. *Discov. Med.* **2010**, *10*, 61–70.
20. Newell, D.G.; Koopmans, M.; Verhoef, L.; Duizer, E.; Aidara-Kane, A.; Sprong, H.; Opsteegh, M.; Langelaar, M.; Threlfall, J.; Scheutz, F.; et al. Food-Borne Diseases—The Challenges of 20 Years Ago Still Persist While New Ones Continue to Emerge. *Int. J. Food Microbiol.* **2010**, *139*, S3–S15. [\[CrossRef\]](#)
21. Martínez-Puchol, S.; Rusiñol, M.; Fernández-Cassi, X.; Timoneda, N.; Itarte, M.; Andrés, C.; Antón, A.; Abril, J.F.; Girones, R.; Bofill-Mas, S. Characterisation of the Sewage Virome: Comparison of NGS Tools and Occurrence of Significant Pathogens. *Sci. Total Environ.* **2020**, *713*, 136604. [\[CrossRef\]](#)
22. Rusiñol, M.; Martínez-Puchol, S.; Timoneda, N.; Fernández-Cassi, X.; Pérez-Cataluña, A.; Fernández-Bravo, A.; Moreno-Mesonero, L.; Moreno, Y.; Alonso, J.L.; Figueras, M.J.; et al. Metagenomic Analysis of Viruses, Bacteria and Protozoa in Irrigation Water. *Int. J. Hyg. Environ. Health* **2020**, *224*, 113440. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Iaconelli, M.; Petricca, S.; Della Libera, S.; Di Bonito, P.; La Rosa, G. First Detection of Human Papillomaviruses and Human Polyomaviruses in River Waters in Italy. *Food Environ. Virol.* **2015**, *7*, 309–315. [\[CrossRef\]](#)
24. La Rosa, G.; Fratini, M.; Accardi, L.; D'Oro, G.; Della Libera, S.; Muscillo, M.; Di Bonito, P. Mucosal and Cutaneous Human Papillomaviruses Detected in Raw Sewages. *PLoS ONE* **2013**, *8*, e52391. [\[CrossRef\]](#)
25. Di Bonito, P.; Della Libera, S.; Petricca, S.; Iaconelli, M.; Sanguinetti, M.; Graffeo, R.; Accardi, L.; La Rosa, G. A Large Spectrum of Alpha and Beta Papillomaviruses Are Detected in Human Stool Samples. *J. Gen. Virol.* **2015**, *96*, 607–613. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Di Cola, G.; Fantilli, A.C.; Pisano, M.B.; Ré, V.E. Foodborne Transmission of Hepatitis A and Hepatitis E Viruses: A Literature Review. *Int. J. Food Microbiol.* **2021**, *338*, 108986. [\[CrossRef\]](#) [\[PubMed\]](#)
27. EFSA BIOHAZ Panel. Scientific Opinion on the Guidance on the Requirements for the Development of Microbiological Criteria. *EFSA J.* **2017**, *15*, 5052. [\[CrossRef\]](#)
28. EU C163. Guidance Document on Addressing Microbiological Risks in Fresh Fruits and Vegetables at Primary Production through Good Hygiene. *Off. J. Eur. Union.* **2017**, *163*, 1–40.
29. Rusiñol, M.; Hundesa, A.; Cárdenas-Youngs, Y.; Fernández-Bravo, A.; Pérez-Cataluña, A.; Moreno-Mesonero, L.; Moreno, Y.; Calvo, M.; Alonso, J.L.; Figueras, M.J.; et al. Microbiological Contamination of Conventional and Reclaimed Irrigation Water: Evaluation and Management Measures. *Sci. Total Environ.* **2020**, *710*, 136298. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Girones, R.; Ferrús, M.A.; Alonso, J.L.; Rodríguez-Manzano, J.; Calgua, B.; de Abreu Corrêa, A.; Hundesa, A.; Carratala, A.; Bofill-Mas, S. Molecular Detection of Pathogens in Water—The Pros and Cons of Molecular Techniques. *Water Res.* **2010**, *44*, 4325–4339. [\[CrossRef\]](#)
31. Gerba, C.P.; Goyal, S.M.; LaBelle, R.L.; Bodgan, G.F. Failure of Indicator Bacteria to Reflect the Occurrence of Enteroviruses in Marine Waters. *Am. J. Public Health* **1979**, *69*, 1116–1119. [\[CrossRef\]](#)
32. Pusch, D.; Oh, D.Y.; Wolf, S.; Dumke, R.; Schröter-Bobsin, U.; Höhne, M.; Röske, I.; Schreier, E. Detection of Enteric Viruses and Bacterial Indicators in German Environmental Waters. *Arch. Virol.* **2005**, *150*, 929–947. [\[CrossRef\]](#) [\[PubMed\]](#)

33. Savichtcheva, O.; Okabe, S. Alternative Indicators of Fecal Pollution: Relations with Pathogens and Conventional Indicators, Current Methodologies for Direct Pathogen Monitoring and Future Application Perspectives. *Water Res.* **2006**, *40*, 2463–2476. [CrossRef] [PubMed]
34. Allard, A.; Vantarakis, A. Adenoviruses. *Glob. Water Pathog. Proj.* **2017**. [CrossRef]
35. Verani, M.; Federigi, I.; Donzelli, G.; Cioni, L.; Carducci, A. Human Adenoviruses as Waterborne Index Pathogens and Their Use for Quantitative Microbial Risk Assessment. *Sci. Total Environ.* **2019**, *651*, 1469–1475. [CrossRef] [PubMed]
36. Hewitt, J.; Greening, G.E.; Leonard, M.; Lewis, G.D. Evaluation of Human Adenovirus and Human Polyomavirus as Indicators of Human Sewage Contamination in the Aquatic Environment. *Water Res.* **2013**, *47*, 6750–6761. [CrossRef]
37. Albinana-Gimenez, N.; Miagostovich, M.P.; Calgua, B.; Huguet, J.M.; Matia, L.; Girones, R. Analysis of Adenoviruses and Polyomaviruses Quantified by QPCR as Indicators of Water Quality in Source and Drinking-Water Treatment Plants. *Water Res.* **2009**, *43*, 2011–2019. [CrossRef]
38. International Organization for Standardization. ISO 15216-1. Microbiology of the Food Chain—Horizontal Method for Determination of Hepatitis A Virus and Norovirus in Food Using Real-Time RT-PCR—Part 1: Method for Quantification. Available online: <https://www.iso.org/standard/65681.html> (accessed on 8 April 2021).
39. Bartsch, C.; Höper, D.; Mäde, D.; John, R. Analysis of Frozen Strawberries Involved in a Large Norovirus Gastroenteritis Outbreak Using Next Generation Sequencing and Digital PCR. *J. Food Microbiol.* **2018**, *76*, 390–395. [CrossRef] [PubMed]
40. Mäde, D.; Trübner, K.; Neubert, E.; Höhne, M.; John, R. Detection and Typing of Norovirus from Frozen Strawberries Involved in a Large-Scale Gastroenteritis Outbreak in Germany. *Food Environ. Virol.* **2013**, *5*, 162–168. [CrossRef] [PubMed]
41. Schrader, C.; Schielke, A.; Ellerbroek, L.; John, R. PCR Inhibitors—Occurrence, Properties and Removal. *J. Appl. Microbiol.* **2012**, *113*, 1014–1026. [CrossRef]
42. Scherer, K.; John, R.; Schrader, C.; Ellerbroek, L.; Schulenburg, J.; Klein, G. Comparison of Two Extraction Methods for Viruses in Food and Application in a Norovirus Gastroenteritis Outbreak. *J. Virol. Methods* **2010**, *169*, 22–27. [CrossRef]
43. Aw, T.G.; Wengert, S.; Rose, J.B. Metagenomic Analysis of Viruses Associated with Field-Grown and Retail Lettuce Identifies Human and Animal Viruses. *Int. J. Food Microbiol.* **2016**, *223*, 50–56. [CrossRef] [PubMed]
44. Park, E.J.; Kim, K.H.; Abell GC, J.; Kim, M.S.; Roh, S.W.; Bae, J.W. Metagenomic Analysis of the Viral Communities in Fermented Foods. *Appl. Environ. Microbiol.* **2011**, *77*, 1284–1291. [CrossRef] [PubMed]
45. Zhang, W.; Li, L.; Deng, X.; Kapusinszky, B.; Delwart, E. What Is for Dinner? Viral Metagenomics of US Store Bought Beef, Pork, and Chicken. *Virology* **2014**, *468–470*, 303–310. [CrossRef]
46. Nieuwenhuijse, D.F.; Koopmans, M.P.G. Metagenomic Sequencing for Surveillance of Food- and Waterborne Viral Diseases. *Front. Microbiol.* **2017**, *8*, 230. [CrossRef]
47. Aarestrup, F.M.; Brown, E.W.; Dettler, C.; Gerner-Smidt, P.; Gilmour, M.W.; Harmsen, D.; Hendriksen, R.S.; Hewson, R.; Heymann, D.L.; Johansson, K.; et al. Integrating Genome-Based Informatics to Modernize Global Disease Monitoring, Information Sharing, and Response. *Emerg. Infect. Dis.* **2012**, *18*. [CrossRef]
48. Desdouits, M.; Wacrenier, C.; Ollivier, J.; Schaeffer, J.; Le Guyader, F.S. A Targeted Metagenomics Approach to Study the Diversity of Norovirus GII in Shellfish Implicated in Outbreaks. *Viruses* **2020**, *12*, 978. [CrossRef] [PubMed]
49. Filipa-Silva, A.; Parreira, R.; Martínez-Puchol, S.; Bofill-Mas, S.; Barreto Crespo, M.T.; Nunes, M. The Unexplored Virome of Two Atlantic Coast Fish: Contribution of Next-Generation Sequencing to Fish Virology. *Foods* **2020**, *9*, 1634. [CrossRef] [PubMed]
50. Strubbia, S.; Schaeffer, J.; Besnard, A.; Wacrenier, C.; Le Mennec, C.; Garry, P.; Desdouits, M.; Le Guyader, F.S. Metagenomic to Evaluate Norovirus Genomic Diversity in Oysters: Impact on Hexamer Selection and Targeted Capture-Based Enrichment. *Int. J. Food Microbiol.* **2020**, *323*, 108588. [CrossRef]
51. Rosani, U.; Shapiro, M.; Venier, P.; Allam, B. A Needle in a Haystack: Tracing Bivalve-Associated Viruses in High-Throughput Transcriptomic Data. *Viruses* **2019**, *11*, 205. [CrossRef]
52. Oshiki, M.; Miura, T.; Kazama, S.; Segawa, T.; Ishii, S.; Hatamoto, M.; Yamaguchi, T.; Kubota, K.; Iguchi, A.; Tagawa, T.; et al. Microfluidic PCR Amplification and MiSeq Amplicon Sequencing Techniques for High-Throughput Detection and Genotyping of Human Pathogenic RNA Viruses in Human Feces, Sewage, and Oysters. *Front. Microbiol.* **2018**, *9*, 830. [CrossRef]
53. Imamura, S.; Kanezashi, H.; Goshima, T.; Haruna, M.; Okada, T.; Inagaki, N.; Uema, M.; Noda, M.; Akimoto, K. Next-Generation Sequencing Analysis of the Diversity of Human Noroviruses in Japanese Oysters. *Foodborne Pathog. Dis.* **2017**, *14*, 465–471. [CrossRef] [PubMed]
54. Chiapponi, C.; Pavoni, E.; Bertasi, B.; Baioni, L.; Scaltriti, E.; Chiesa, E.; Cianti, L.; Losio, M.N.; Pongolini, S. Isolation and Genomic Sequence of Hepatitis A Virus from Mixed Frozen Berries in Italy. *Food Environ. Virol.* **2014**, *6*, 202–206. [CrossRef] [PubMed]
55. Yang, Z.; Mammel, M.; Papafragkou, E.; Hida, K.; Elkins, C.A.; Kulka, M. Application of Next Generation Sequencing toward Sensitive Detection of Enteric Viruses Isolated from Celery Samples as an Example of Produce. *Int. J. Food Microbiol.* **2017**, *261*, 73–81. [CrossRef] [PubMed]
56. Toscano, A.; Marzo, A.; Milani, M.; Cirelli, G.L.; Barbagallo, S. Comparison of Removal Efficiencies in Mediterranean Pilot Constructed Wetlands Vegetated with Different Plant Species. *Ecol. Eng.* **2015**, *75*, 155–160. [CrossRef]
57. Ribot, M.; Bernal, S.; Nikolakopoulou, M.; Vaessen, T.N.; Cochero, J.; Gacia, E.; Sorolla, A.; Argerich, A.; Sabater, F.; Isnard, M.; et al. Enhancement of Carbon and Nitrogen Removal by Helophytes along Subsurface Water Flowpaths Receiving Treated Wastewater. *Sci. Total Environ.* **2017**, *599–600*, 1667–1676. [CrossRef]

58. Bofill-Mas, S.; Albinana-Gimenez, N.; Clemente-Casares, P.; Hundesa, A.; Rodriguez-Manzano, J.; Allard, A.; Calvo, M.; Girones, R. Quantification and Stability of Human Adenoviruses and Polyomavirus JCPyV in Wastewater Matrices. *Appl. Environ. Microbiol.* **2006**, *72*, 7894–7896. [\[CrossRef\]](#)
59. Da Silva, A.K.; Le Saux, J.C.; Parnaudeau, S.; Pommepuy, M.; Elimelech, M.; Le Guyader, F.S. Evaluation of Removal of Noroviruses during Wastewater Treatment, Using Real-Time Reverse Transcription-PCR: Different Behaviors of Genogroups I and II. *Appl. Environ. Microbiol.* **2007**, *73*, 7891–7897. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Svraaka, S.; Duizer, E.; Vennema, H.; De Bruin, E.; Van Der Veer, B.; Dorresteyn, B.; Koopmans, M. Etiological Role of Viruses in Outbreaks of Acute Gastroenteritis in The Netherlands from 1994 through 2005. *J. Clin. Microbiol.* **2007**, *45*, 1389–1394. [\[CrossRef\]](#)
61. Hoehne, M.; Schreier, E. Detection of Norovirus Genogroup I and II by Multiplex Real-Time RT-PCR Using a 3'-Minor Groove Binder-DNA Probe. *BMC Infect. Dis.* **2006**, *6*, 69. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Loisy, F.; Atmar, R.L.; Guillon, P.; Le Cann, P.; Pommepuy, M.; Le Guyader, F.S. Real-Time RT-PCR for Norovirus Screening in Shellfish. *J. Virol. Methods* **2005**, *123*, 1–7. [\[CrossRef\]](#)
63. Kageyama, T.; Kojima, S.; Shinohara, M.; Uchida, K.; Fukushi, S.; Hoshino, F.B.; Takeda, N.; Katayama, K. Broadly Reactive and Highly Sensitive Assay for Norwalk-like Viruses Based on Real-Time Quantitative Reverse Transcription-PCR. *J. Clin. Microbiol.* **2003**, *41*, 1548–1557. [\[CrossRef\]](#)
64. Jothikumar, N.; Cromeans, T.L.; Robertson, B.H.; Meng, X.J.; Hill, V.R. A Broadly Reactive One-Step Real-Time RT-PCR Assay for Rapid and Sensitive Detection of Hepatitis E Virus. *J. Virol. Methods* **2006**, *131*, 65–71. [\[CrossRef\]](#)
65. Fernandez-Cassi, X.; Timoneda, N.; Martínez-Puchol, S.; Rusiñol, M.; Rodriguez-Manzano, J.; Figuerola, N.; Bofill-Mas, S.; Abril, J.F.; Girones, R. Metagenomics for the Study of Viruses in Urban Sewage as a Tool for Public Health Surveillance. *Sci. Total. Environ.* **2018**, *618*, 870–880. [\[CrossRef\]](#)
66. Fernandez-Cassi, X.; Rusiñol, M.; Martínez-Puchol, S. Viral Concentration and Amplification from Human Serum Samples Prior to Application of Next-Generation Sequencing Analysis. *Methods Mol. Biol.* **2018**, *1838*, 173–188. [\[CrossRef\]](#) [\[PubMed\]](#)
67. Hjelmsø, M.H.; Møllerup, S.; Jensen, R.H.; Pietroni, C.; Lukjancenko, O.; Schultz, A.C.; Aarestrup, F.M.; Hansen, A.J. Metagenomic Analysis of Viruses in Toilet Waste from Long Distance Flights—A New Procedure for Global Infectious Disease Surveillance. *PLoS ONE* **2019**, *14*, e0210368. [\[CrossRef\]](#)
68. Briese, T.; Kapoor, A.; Mishra, N.; Jain, K.; Kumar, A.; Jabado, O.J.; Ian Lipkina, W. Virome Capture Sequencing Enables Sensitive Viral Diagnosis and Comprehensive Virome Analysis. *MBio* **2015**, *6*, e01491-15. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Vilsker, M.; Moosa, Y.; Nooij, S.; Fonseca, V.; Ghysens, Y.; Dumon, K.; Pauwels, R.; Alcantara, L.C.; Vanden Eynden, E.; Vandamme, A.-M.; et al. Genome Detective: An Automated System for Virus Identification from High-Throughput Sequencing Data. *Bioinformatics* **2019**, *35*, 871–873. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Benson, D.A.; Clark, K.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Sayers, E.W. GenBank. *Nucleic Acids Res.* **2015**, *43*, D30–D35. [\[CrossRef\]](#)
71. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [\[CrossRef\]](#)
72. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An Integrated and Extendable Desktop Software Platform for the Organization and Analysis of Sequence Data. *Bioinformatics* **2012**, *28*, 1647–1649. [\[CrossRef\]](#)
73. Chhabra, P.; de Graaf, M.; Parra, G.I.; Chan, M.C.W.; Green, K.; Martella, V.; Wang, Q.; White, P.A.; Katayama, K.; Vennema, H.; et al. Updated Classification of Norovirus Genogroups and Genotypes. *J. Gen. Virol.* **2019**, *100*, 1393–1406. [\[CrossRef\]](#)
74. Chhabra, P.; de Graaf, M.; Parra, G.I.; Chan, M.C.W.; Green, K.; Martella, V.; Wang, Q.; White, P.A.; Katayama, K.; Vennema, H.; et al. Corrigendum: Updated Classification of Norovirus Genogroups and Genotypes. *J. Gen. Virol.* **2020**, *101*, 893. [\[CrossRef\]](#)
75. Kroneman, A.; Vennema, H.; Deforche, K.; Avoort, H.; Peñaranda, S.; Oberste, M.S.; Vinjé, J.; Koopmans, M. An Automated Genotyping Tool for Enteroviruses and Noroviruses. *J. Clin. Virol.* **2011**, *51*, 121–125. [\[CrossRef\]](#)
76. Kitajima, M.; Oka, T.; Haramoto, E.; Takeda, N.; Katayama, K.; Katayama, H. Seasonal Distribution and Genetic Diversity of Genogroups I, II, and IV Noroviruses in the Tamagawa River, Japan. *Environ. Sci. Technol.* **2010**, *44*, 7116–7122. [\[CrossRef\]](#)
77. Aw, T.G.; Gin, K.Y.H.; Oon, L.L.E.; Chen, E.X.; Woo, C.H. Prevalence and Genotypes of Human Noroviruses in Tropical Urban Surface Waters and Clinical Samples in Singapore. *Appl. Environ. Microbiol.* **2009**, *75*, 4984–4992. [\[CrossRef\]](#)
78. Forslund, O.; Ly, H.; Higgins, G. Improved Detection of Cutaneous Human Papillomavirus DNA by Single Tube Nested “hanging Droplet” PCR. *J. Virol. Methods* **2003**, *110*, 129–136. [\[CrossRef\]](#)
79. De Roda Husman, A.M.; Walboom, J.M.M.; Van den Brule, A.J.C.; Meijer, C.J.L.M.; Snijders, P.J.F. The Use of General Primers GP5 and GP6 Elongated at Their 3' Ends with Adjacent Highly Conserved Sequences Improves Human Papillomavirus Detection by PCR. *J. Gen. Virol.* **1995**, *76*, 1057–1062. [\[CrossRef\]](#) [\[PubMed\]](#)
80. Bofill-Mas, S.; Rusiñol, M.; Fernandez-Cassi, X.; Carratalà, A.; Hundesa, A.; Girones, R. Quantification of Human and Animal Viruses to Differentiate the Origin of the Fecal Contamination Present in Environmental Samples. *Biomed. Res. Int.* **2013**, *2013*. [\[CrossRef\]](#) [\[PubMed\]](#)
81. Rusiñol, M.; Fernandez-Cassi, X.; Hundesa, A.; Vieira, C.; Kern, A.; Eriksson, I.; Ziros, P.; Kay, D.; Miagostovich, M.; Vargha, M.; et al. Application of Human and Animal Viral Microbial Source Tracking Tools in Fresh and Marine Waters from Five Different Geographical Areas. *Water Res.* **2014**, *59*, 119–129. [\[CrossRef\]](#)
82. Rusiñol, M.; Girones, R. Summary of Excreted and Waterborne Viruses. *Glob. Water Pathog. Proj.* **2017**. [\[CrossRef\]](#)

83. Vieira, C.B.; de Abreu Corrêa, A.; de Jesus, M.S.; Luz, S.L.B.; Wyn-Jones, P.; Kay, D.; Vargha, M.; Miagostovich, M.P. Viruses Surveillance Under Different Season Scenarios of the Negro River Basin, Amazonia, Brazil. *Food Environ. Virol.* **2016**, *8*, 57–69. [\[CrossRef\]](#) [\[PubMed\]](#)
84. Kokkinos, P.; Kozyra, I.; Lazic, S.; Bouwknegt, M.; Rutjes, S.; Willems, K.; Moloney, R.; de Roda Husman, A.M.; Kaupke, A.; Legaki, E.; et al. Harmonised Investigation of the Occurrence of Human Enteric Viruses in the Leafy Green Vegetable Supply Chain in Three European Countries. *Food Environ. Virol.* **2012**, *4*, 179–191. [\[CrossRef\]](#)
85. Maunula, L.; Kaupke, A.; Vasickova, P.; Söderberg, K.; Kozyra, I.; Lazic, S.; van der Poel, W.H.M.; Bouwknegt, M.; Rutjes, S.; Willems, K.A.; et al. Tracing Enteric Viruses in the European Berry Fruit Supply Chain. *Int. J. Food Microbiol.* **2013**, *167*, 177–185. [\[CrossRef\]](#) [\[PubMed\]](#)
86. Brassard, J.; Gagné, M.; Gagné, M.; Côté, C. Detection of Human Food-Borne and Zoonotic Viruses on Irrigated, Field-Grown Strawberries. *Appl. Environ. Microbiol.* **2012**, *78*, 3763–3766. [\[CrossRef\]](#) [\[PubMed\]](#)
87. Baert, L.; Mattison, K.; Loisy-Hamon, F.; Harlow, J.; Martyres, A.; Lebeau, B.; Stals, A.; Van Coillie, E.; Herman, L.; Uyttendaele, M. Review: Norovirus Prevalence in Belgian, Canadian and French Fresh Produce: A Threat to Human Health? *Int. J. Food Microbiol.* **2011**, *151*, 261–269. [\[CrossRef\]](#)
88. Cantalupo, P.G.; Calgua, B.; Zhao, G.; Hundesa, A.; Wier, A.D.; Katz, J.P.; Grabe, M.; Hendrix, R.W.; Girones, R.; Wang, D.; et al. Raw Sewage Harbors Diverse Viral Populations. *MBio* **2011**, *2*, e00180. [\[CrossRef\]](#)
89. De Graaf, M.; Van Beek, J.; Vennema, H.; Podkolzin, A.T.; Hewitt, J.; Bucardo, F.; Templeton, K.; Mans, J.; Nordgren, J.; Reuter, G.; et al. Emergence of a Novel GII.17 Norovirus—End of the GII.4 Era? *Eurosurveillance* **2015**, *20*, 21178. [\[CrossRef\]](#)
90. Le Guyader, F.S.; Mittelholzer, C.; Haugarrau, L.; Hedlund, K.O.; Alsterlund, R.; Pommepuy, M.; Svensson, L. Detection of Noroviruses in Raspberries Associated with a Gastroenteritis Outbreak. *Int. J. Food Microbiol.* **2004**, *97*, 179–186. [\[CrossRef\]](#)
91. Ogorzal, L.; Walczak, C.; Galloux, M.; Etienne, S.; Gassilloud, B.; Cauchie, H.M. Human Adenovirus Diversity in Water Samples Using a Next-Generation Amplicon Sequencing Approach. *Food Environ. Virol.* **2015**, *7*, 112–121. [\[CrossRef\]](#)
92. Hamza, H.; Hamza, I.A. Oncogenic Papillomavirus and Polyomavirus in Urban Sewage in Egypt. *Sci. Total Environ.* **2018**, *610–611*, 1413–1420. [\[CrossRef\]](#)
93. La Rosa, G.; Della Libera, S.; Petricca, S.; Iaconelli, M.; Briancesco, R.; Paradiso, R.; Semproni, M.; di Bonito, P.; Bonadonna, L. First Detection of Papillomaviruses and Polyomaviruses in Swimming Pool Waters: Unrecognized Recreational Water-Related Pathogens? *J. Appl. Microbiol.* **2015**, *119*, 1683–1691. [\[CrossRef\]](#)
94. Houde, A.; Leblanc, D.; Poitras, E.; Ward, P.; Brassard, J.; Simard, C.; Trottier, Y.L. Comparative Evaluation of RT-PCR, Nucleic Acid Sequence-Based Amplification (NASBA) and Real-Time RT-PCR for Detection of Noroviruses in Fecal Material. *J. Virol. Methods* **2006**, *135*, 163–172. [\[CrossRef\]](#) [\[PubMed\]](#)
95. Fernandez-Cassi, X.; Martínez-Puchol, S.; Silva-Sales, M.; Cornejo, T.; Bartolome, R.; Bofill-Mas, S.; Girones, R. Unveiling Viruses Associated with Gastroenteritis Using a Metagenomics Approach. *Viruses* **2020**, *12*, 1432. [\[CrossRef\]](#) [\[PubMed\]](#)
96. Tohma, K.; Lepore, C.J.; Gao, Y.; Ford-Siltz, L.A.; Parra, G.I. Population Genomics of GII.4 Noroviruses Reveal Complex Diversification and New Antigenic Sites Involved in the Emergence of Pandemic Strains. *MBio* **2019**, *10*, e02202-19. [\[CrossRef\]](#)
97. Maunula, L.; Roivainen, M.; Keränen, M.; Mäkelä, S.; Söderberg, K.; Summa, M.; von Bonsdorff, C.H.; Lappalainen, M.; Korhonen, T.; Kuusi, M.; et al. Detection of Human Norovirus from Frozen Raspberries in a Cluster of Gastroenteritis Outbreaks. *Eurosurveillance* **2009**, *14*, 19435. [\[CrossRef\]](#) [\[PubMed\]](#)
98. Jeong, H.S.; Jeong, A.; Cheon, D.S. Epidemiology of Astrovirus Infection in Children. *Korean J. Pediatr.* **2012**, *55*, 77–82. [\[CrossRef\]](#) [\[PubMed\]](#)
99. Johnson, C.; Hargest, V.; Cortez, V.; Meliopoulos, V.A.; Schultz-Cherry, S. Astrovirus Pathogenesis. *Viruses* **2017**, *9*, 22. [\[CrossRef\]](#)
100. Cordey, S.; Vu, D.L.; Schibler, M.; L'Huillier, A.G.; Brito, F.; Docquier, M.; Posfay-Barbe, K.M.; Petty, T.J.; Turin, L.; Zdobnov, E.M.; et al. Astrovirus MLB2, a New Gastroenteric Virus Associated with Meningitis and Disseminated Infection. *Emerg. Infect. Dis.* **2016**, *22*, 846–853. [\[CrossRef\]](#)
101. Kapoor, A.; Li, L.; Victoria, J.; Oderinde, B.; Mason, C.; Pandey, P.; Zaidi, S.Z.; Delwart, E. Multiple Novel Astrovirus Species in Human Stool. *J. Gen. Virol.* **2009**, *90*, 2965–2972. [\[CrossRef\]](#)
102. Gabbay, Y.B.; Leite, J.P.G.; Oliveira, D.S.; Nakamura, L.S.; Nunes, M.R.T.; Mascarenhas, J.D.A.P.; Heinemann, M.B.; Linhares, A.C. Molecular Epidemiology of Astrovirus Type 1 in Belém, Brazil, as an Agent of Infantile Gastroenteritis, over a Period of 18 Years (1982–2000): Identification of Two Possible New Lineages. *Virus Res.* **2007**, *129*, 166–174. [\[CrossRef\]](#)
103. Rivadulla, E.; Romalde, J.L. A Comprehensive Review on Human Aichi Virus. *Virol. Sin.* **2020**, *35*, 501–516. [\[CrossRef\]](#) [\[PubMed\]](#)
104. Reuter, G.; Boros, A.; Pankovics, P. Kobuviruses—A Comprehensive Review. *Rev. Med. Virol.* **2011**, *21*, 32–41. [\[CrossRef\]](#)
105. Ambert-Galay, K.; Lorrot, M.; Bon, F.; Giraudon, H.; Kaplon, J.; Wolfer, M.; Lebon, P.; Gendrel, D.; Pothier, P. Prevalence and Genetic Diversity of Aichi Virus Strains in Stool Samples from Community and Hospitalized Patients. *J. Clin. Microbiol.* **2008**, *46*, 1252–1258. [\[CrossRef\]](#)
106. Rosario, K.; Nilsson, C.; Lim, Y.W.; Ruan, Y.; Breitbart, M. Metagenomic Analysis of Viruses in Reclaimed Water. *Environ. Microbiol.* **2009**, *11*, 2806–2820. [\[CrossRef\]](#) [\[PubMed\]](#)
107. Gerba, C.P.; Choi, C.Y. Role of Irrigation Water in Crop Contamination by Viruses. In *Viruses in Foods*; Springer: Boston, MA, USA, 2006; pp. 257–263. [\[CrossRef\]](#)
108. Wei, J.; Jin, Y.; Sims, T.; Kniel, K.E. Internalization of Murine Norovirus 1 by *Lactuca Sativa* during Irrigation. *Appl. Environ. Microbiol.* **2011**, *77*, 2508–2512. [\[CrossRef\]](#) [\[PubMed\]](#)

109. Reganold, J.P.; Wachter, J.M. Organic Agriculture in the Twenty-First Century. *Nat. Plants* **2016**, *2*, 15221. [[CrossRef](#)]
110. Machado, D.C.; Maia, C.M.; Carvalho, I.D.; Da Silva, N.F.; Dantas Porfírio Borges Andre, M.C.; Serafini, Á.B. Microbiological Quality of Organic Vegetables Produced in Soil Treated with Different Types of Manure and Mineral Fertilizer. *Brazilian J. Microbiol.* **2006**, *37*, 538–544. [[CrossRef](#)]
111. Grusenmeyer, D.C.; Cramer, T.N. Symposium: Manure Management. *J. Dairy Sci.* **1997**, *80*, 2651–2654. [[CrossRef](#)]
112. Pell, A.N. Manure and Microbes: Public and Animal Health Problem? *J. Dairy Sci.* **1997**, *80*, 2673–2681. [[CrossRef](#)]
113. Alegbeleye, O.O.; Singleton, I.; Sant’Ana, A.S. Sources and Contamination Routes of Microbial Pathogens to Fresh Produce during Field Cultivation: A Review. *Food Microbiol.* **2018**, *73*, 177–208. [[CrossRef](#)]
114. Mawdsley, J.L.; Bardgett, R.D.; Merry, R.J.; Pain, B.F.; Theodorou, M.K. Pathogens in Livestock Waste, Their Potential for Movement through Soil and Environmental Pollution. *Appl. Soil Ecol.* **1995**, *2*, 1–15. [[CrossRef](#)]
115. Duhaime, M.B.; Deng, L.; Poulos, B.T.; Sullivan, M.B. Towards Quantitative Metagenomics of Wild Viruses and Other Ultra-Low Concentration DNA Samples: A Rigorous Assessment and Optimization of the Linker Amplification Method. *Environ. Microbiol.* **2012**, *14*, 2526–2537. [[CrossRef](#)] [[PubMed](#)]
116. Karlsson, O.E.; Belák, S.; Granberg, F. The Effect of Preprocessing by Sequence-Independent, Single-Primer Amplification (SISPA) on Metagenomic Detection of Viruses. *Biosecur. Bioterror. Biodef. Strategy Pract. Sci.* **2013**, *11*, S227–S234. [[CrossRef](#)]
117. Li, J.; Pan, Y.Q.; Xu, Z.; Wang, Q.; Hang, D.; Shen, N.; Liu, M.; Zhang, C.; Abliz, A.; Deng, Q.; et al. Improved Detection of Human Papillomavirus Harbored in Healthy Skin with FAP6085/64 Primers. *J. Virol. Methods* **2013**, *193*, 633–638. [[CrossRef](#)] [[PubMed](#)]
118. Akgül, B.; Cooke, J.C.; Storey, A. HPV-Associated Skin Disease. *J. Pathol.* **2006**, *208*, 165–175. [[CrossRef](#)] [[PubMed](#)]
119. Hošnjak, L.; Komloš, K.F.; Kocjan, B.J.; Seme, K.; Poljak, M. Development of a Novel Multiplex Type-Specific Quantitative Realtime PCR for Detection and Differentiation of Infections with Human Papillomavirus Types HPV2, HPV27, and HPV57. *Acta Dermatovenerol. APA* **2016**, *25*, 65–71. [[CrossRef](#)]
120. Yeni, F.; Alpas, H. Vulnerability of Global Food Production to Extreme Climatic Events. *Food Res. Int.* **2017**, *96*, 27–39. [[CrossRef](#)]
121. Berger, C.N.; Sodha, S.V.; Shaw, R.K.; Griffin, P.M.; Pink, D.; Hand, P.; Frankel, G. Fresh Fruit and Vegetables as Vehicles for the Transmission of Human Pathogens. *Environ. Microbiol.* **2010**, *12*, 2385–2397. [[CrossRef](#)]
122. Kokkinos, P.; Kozyra, I.; Lazic, S.; Söderberg, K.; Vasickova, P.; Bouwknegt, M.; Rutjes, S.; Willems, K.; Moloney, R.; de Roda Husman, A.M.; et al. Virological quality of irrigation water in leafy green vegetables and berry fruits production chains. *Food Environ. Virol.* **2017**, *9*, 72–78. [[CrossRef](#)]

Article 5: Explorant la contaminació vírica en aigües subterrànies i d'escorrentia en un entorn urbà

"Exploring Viral Contamination in Urban Groundwater and Runoff"

Itarte, M., Forés, E., Martínez-Puchol, S., Scheiber, L., Vázquez-Suñé, E., Bofill-Mas, S. i Rusiñol, M.

Article en procés de revisió a la revista Science of the Total Environment.

La dependència dels aqüífers urbans per part de la població mundial està augmentant de manera constant, i la monitorització de la qualitat de les aigües subterrànies és insuficient. Els aqüífers urbans són susceptibles a la contaminació de patògens a través de diverses fonts de contaminació, com ara les fuites de clavegueres o les escorrenties urbanes.

En aquest estudi, es va analitzar la presència d'indicadors fecals i patògens vírics en aigües residuals, aigües subterrànies i escorrenties urbanes mitjançant assaigs de qPCR. A més, es va aplicar la seqüenciació massiva amb enriquiment de dianes mitjançant el panell de captura VirCapSeq-VERT per explorar la diversitat de virus de vertebrats present en les mostres d'aigües subterrànies i d'escorrentia, oferint informació sobre la contaminació vírica i les possibles rutes de transmissió de virus en àrees urbanes.

Es va identificar HAdV en totes les mostres d'aigües residuals, en el 67% de les mostres d'aigua subterrània i en una mostra d'escorrentia mitjançant qPCR, indicant contaminació fecal humana. NoV del genogrup I (NoV GI) es va detectar en aigües residuals i en dues mostres d'aigua subterrània d'hivern de zones altament i mitjanament urbanitzades. NoV del genogrup II (NoV GII), EV i SARS-CoV-2 es van detectar

exclusivament en aigües residuals. La seqüenciació massiva utilitzant el panell VirCapSeq-VERT va permetre la detecció de virus humans i animals en aigües subterrànies i d'escorrentia. En aigües subterrànies, es va identificar HPV, així com circovirus i picornavirus animals que infecten ratpenats, coloms o gossos. D'altra banda, les mostres d'escorrentia van mostrar una major diversitat vírica, incloent-hi virus que infecten humans, coloms, animals domèstics o ratpenats, i que pertanyen a les famílies víriques *Anelloviridae*, *Astroviridae*, *Caliciviridae*, *Circoviridae*, *Coronaviridae*, *Genomoviridae*, *Parvoviridae*, *Picobirnaviridae*, *Picornaviridae*, *Poxviridae* i *Retroviridae*.

Aquest estudi ofereix informació sobre el viroma present en aigües subterrànies i d'escorrentia urbanes. Les aigües subterrànies urbanes s'utilitzen sovint per al reg de parcs i jardins, la neteja de carrers, etc. No obstant, en els últims anys, s'ha plantejat l'objectiu d'ampliar les seves aplicacions, considerant la possibilitat d'utilitzar-les per al subministrament d'aigua de la població. Així doncs, amb la creixent demanda de recursos hídrics, la contaminació vírica d'aquestes aigües podria resultar en un augment de la incidència de malalties transmeses per l'aigua. La identificació de virus patògens presents en les aigües subterrànies mitjançant la seqüenciació massiva amb enriquiment de dianes és important per avaluar els riscos potencials associats amb les malalties transmeses per l'aigua, especialment en àrees urbanes cada vegada més poblades però també en àrees rurals que depenen de l'ús d'aigües subterrànies. D'entre les possibles fonts de contaminació, les aigües d'escorrentia urbanes poden representar un risc important si contenen virus patògens. Per tant, identificar les fonts de contaminació i monitoritzar els virus humans i animals amb potencial zoonòtic en aigües subterrànies i d'escorrentia, ajuda a avaluar els riscos per a la salut pública i permet desenvolupar mesures de mitigació efectives.

Exploring Viral Contamination in Urban Groundwater and Runoff

Marta Itarte^{1,2,}, Eva Forés^{1,2}, Sandra Martínez-Puchol^{1,3}, Laura Scheiber⁴,
Enric Vázquez-Suñé⁴, Sílvia Bofill-Mas^{1,2} and Marta Rusiñol^{1,2}*

¹Laboratory of Viruses Contaminants of Water and Food, Secció de Microbiologia, Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona (UB), 08028 Barcelona, Spain

²Institut de Recerca de l'Aigua (IdRA), Universitat de Barcelona (UB), 08001 Barcelona, Spain

³Vicerectorat de Recerca, Universitat de Barcelona (UB), 08007 Barcelona, Spain

⁴Department of Geosciences, Institute of Environmental Assessment and Water Research (IDAEA), Severo Ochoa Excellence Center of the Spanish Council for Scientific Research (CSIC), 08034 Barcelona, Spain

*Corresponding author: mitarte@ub.edu

KEYWORDS

Groundwater virome; runoff virome; next-generation sequencing; virus contamination sources.

ABSTRACT

The reliance of the global population on urban aquifers is steadily increasing, and there is insufficient monitoring of groundwater quality. Urban aquifers are susceptible to pathogenic contamination through sources such as sewer leakage or urban runoff. In this study, quantitative polymerase chain reaction (qPCR) was employed to evaluate the presence of human fecal viral indicators and viral pathogens in urban wastewater, groundwater, and runoff. Additionally, a target enrichment sequencing (TES) approach was utilized to explore the viral diversity within groundwater and runoff samples, offering insights into viral contamination and potential virus transmission routes in urban areas.

Human adenovirus (HAdV) was identified in all wastewater samples, 67% of groundwater samples, and one runoff sample by qPCR indicating human viral fecal contamination. The viral pathogen NoV GI was detected in wastewater and two winter groundwater samples from highly and medium urbanized areas. NoV GII, EV and SARS-CoV-2 were exclusively detected in wastewater. Human and other vertebrate viruses were detected in groundwater and runoff samples using TES. This study gives insights about the virome present in urban water sources, emphasizing the need for thorough monitoring and deeper understanding to address emerging public health concerns.

1. INTRODUCTION

The proportion of the global population relying on urban aquifers as drinking water sources is growing annually, especially during periods of drought. However, in most cases, there is insufficient monitoring of groundwater quality. Large amounts of pathogens may be introduced into urban aquifers through multiple sources and pathways, mainly from sewer leakage but also urban runoff, sewer overflows, and wastewater treatment plant effluent recharge areas (1).

The primary contributing factor to disease outbreaks associated with untreated groundwater appears to be human sewage, and indeed, these outbreaks are more likely to have a viral etiology (2). In highly urbanized areas, rainfall events generate highly polluted stormwater runoff, which has also been identified as a significant source of microbial contamination to receiving bodies of water (3). Human and animal viruses of fecal origin, including SARS-CoV-2, have been identified in stormwater collected from urban areas (4–9), highlighting runoff as a potential transmission pathway (10).

The study of groundwater virome is indeed an emerging field, and while there might be limited research (1), it is plausible to assume that the groundwater virome possesses a rich diverse array of viruses. The presence of a high diversity of viruses in both raw wastewater and treated wastewater (11,12) suggests that a similar diversity might exist in groundwater impacted by sewage. Next-generation sequencing (NGS) techniques enable the identification of viruses in environmental samples and can provide a comprehensive overview of the groundwater virome, with plant viruses and bacteriophages being the most widely prevalent viruses in groundwater environments (13). Additionally, target enrichment sequencing (TES) has demonstrated to improve the detection of vertebrate viruses that would otherwise be difficult to identify due to their low concentrations in environmental samples, such as groundwater used for irrigation (14).

This study aimed to explore the viral diversity in both urban groundwater and runoff, and providing insights into viral contamination and potential transmission routes of viruses in urban areas.

2. MATERIALS AND METHODS

2.1 Sampling and sample processing

Four locations in the urban area of Barcelona were chosen for groundwater and sewer wastewater sampling, representing areas with different degrees of urbanization: highly urbanized (26,288 hab/km²), medium urbanized (7,590 hab/km²), industrial district (8,272 hab/km²), and riverside areas (23,271 hab/km²) (15). One-hundred liters of groundwater and grab wastewater samples (80 mL) from the nearest sewer system were collected in parallel over two sampling campaigns. A total of 7 groundwater and 8 wastewater samples were collected in summer (June and July 2020) and winter (December 2020) seasons, with two sampling events per season. Industrial district wastewater sample was not collected during the winter campaign due to intensive rainy events and restricted access to the sewer system. Two sites from the medium urbanized district

were selected for runoff sampling and up to 8 L were collected for each site (RO1 and RO2) after storm events (17.2 mm of accumulated precipitation) occurred in November 2020.

Viruses were concentrated from groundwater and runoff samples by ultrafiltration using the Long Volume Concentration kit (LVC kit) from InnovaPrep® (InnovaPrep, Drexel, MO, USA) that uses a Rexeed 25-A polysulfone hollow-fiber ultrafilter and wet foam elution cans with Tris-PBS. The eluted volume (50-100 mL) was subjected to a secondary concentration step with Centricon® Plus-70 devices with a cut-off of 30 kDa (Merck Millipore, Burlington, MA, USA), into a final eluted volume of 200-300 µL (16). Wastewater samples were concentrated directly with Centricon® Plus-70 as described before (17). The nucleic acids were extracted using the QIAamp® Viral RNA Mini Kit from QIAGEN (QIAGEN, Germantown, MD, USA) into 100 µl and stored at -80°C for further analysis in quantitative polymerase chain reaction (qPCR) and NGS assays.

2.2 Viral quantification

Viral contamination was assessed by specific qPCR assays for HAdV using TaqMan® Environmental Master Mix 2.0 (Applied Biosystems, Waltham, MA, USA) with the specific primers and probe previously described (18,19). Viral RNA pathogens were also quantified in all samples using the RNA UltraSense™ One-Step qRT-PCR System (Applied Biosystems, Waltham, MA, USA) with specific primers and probes for NoV GI (20–22), NoV GII (23,24), EV (25) and SARS-CoV-2 (26). With the exception of the synthetic SARS-CoV-2 control (control 51 from Twist Biosciences), the qPCR standards were prepared using synthetic gBlocks® Gene fragments (IDT, Coralville, IA, USA). All qPCR assays were performed using the QuantStudio™ Real-Time PCR System from ThermoFisher Scientific. Both undiluted and 10-fold diluted nucleic acid extractions were analyzed, and non-template controls were included in each of the assays.

2.3 Metagenomic analysis: target enrichment sequencing (TES)

To explore the diversity of vertebrate viral pathogens present in groundwater and runoff samples, nucleic acid extractions were analyzed using the VirCapSeqVERT Capture Panel (Roche, Basel, Switzerland). This panel employs probes designed for the capture of sequences from vertebrate viral pathogens through hybridization, allowing the detection of these sequences in complex sample types, as demonstrated in previous studies (11,27–33).

A sequence-independent, single-primer amplification (SISPA) method was performed, and the KAPA HyperPlus Library Preparation Kit (KAPA Biosystems, Roche, Basel, Switzerland) along with the KAPA UDI Primer Mixes (KAPA Biosystems, Roche, Basel, Switzerland) were used to prepare the sequencing libraries as described in previous studies (30). The libraries were equimolarly pooled, captured using the VirCapSeqVERT Capture Panel and sequenced on an Illumina NextSeq 500 platform using a 2 x 150 cycles Mid Output run.

Bioinformatic analysis was performed using the CZ ID portal (<https://czid.org/>, accessed on 6 February 2024), an open-source cloud-based pipeline and service designed for metagenomic pathogen detection and monitoring (34). Briefly, this portal performs host and quality filtration steps, followed by an assembly-based alignment pipeline, which results in the assignment of reads and contigs into taxonomic categories. Only viral assignments with a nucleotide identity above 70% and an alignment length exceeding 100 bp were considered in the analysis.

3. RESULTS

3.1 Viral quantification in wastewater, groundwater, and runoff samples

Urban groundwater and its sources of viral contamination, wastewater and runoff, were analyzed for viral contaminants by specific qPCR assays. The presence of HAdV, an indicator of human fecal contamination, was identified in wastewater samples across all study locations during both winter and summer seasons with concentrations ranging from 1.18E+03 to 5.18E+06 GC/L. This viral indicator was also detected in 67% of groundwater samples with concentrations ranging from 4.67E+00 to 4.67E+02 GC/100 L. Additionally, it was found in one runoff sample at a concentration of 3.00E+00 GC/L, while no other viruses were identified in runoff samples using qPCR assays.

Figure 1 provides a summary of the mean viral concentrations observed in wastewater and groundwater samples during the summer. HAdV concentration values in summer wastewater samples ranged from 1.18E+03 to 6.92E+04 GC/L. NoV GI and GII were detected in all wastewater locations, with the exception of NoV GII that was not detected in medium urbanized wastewater samples. NoV GI concentrations in wastewater ranged from 4.45E+02 to 1.34E+05 GC/L, whereas NoV GII ranged from 1.13E+04 to 1.02E+05 GC/L. EV was exclusively detected in one riverside wastewater sample, with a concentration of 2.94E+05 GC/L. No viral pathogens, rather than HAdV, were detected in urban groundwater samples collected during the summer season. Groundwater from all locations, except in highly urbanized areas, exhibited low levels of contamination with HAdV, up to 2.25E+02 GC/100 L.

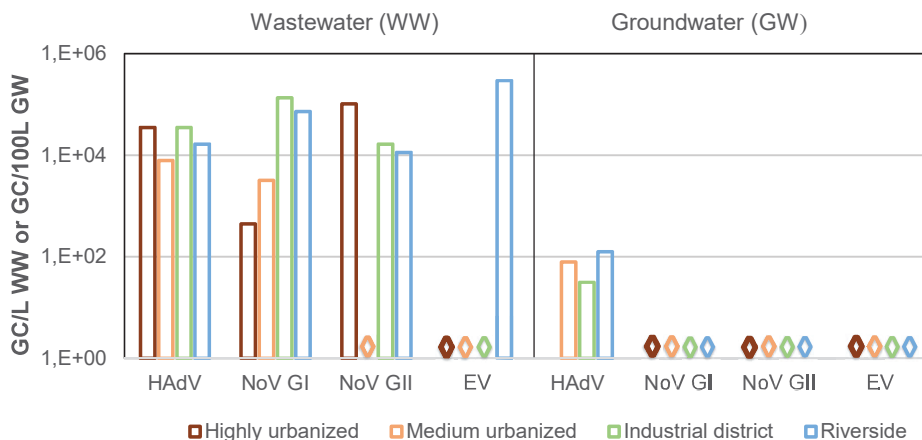


Figure 1. Viral pathogens detected and quantified in wastewater (WW) and groundwater (GW) samples from the summer campaign. Bars represent mean values of the concentrations obtained by qPCR assays. Samples that tested negative are indicated with $\Diamond$.

The mean viral concentrations detected in winter wastewater and groundwater samples are presented in Figure 2. As for the summer campaign, HAdV was identified in all wastewater samples, with concentrations ranging from $3.50E+04$ to $5.18E+06$ GC/L. NoVs were detected in wastewater from highly urbanized and riverside district areas, with concentrations ranging from $5.89E+01$ to $7.50E+05$ GC/L for NoV GI, and from $1.35E+04$ to $3.12E+06$ GC/L for NoV GII. EV was not found in any wastewater samples. Contamination by both HAdV and NoV GI was observed in highly urbanized and medium urbanized groundwater, whereas industrial district and riverside groundwater showed only HAdV contamination. The highest concentration value of HAdV, $4.67E+02$ GC/100L, was observed in a sample from the highly urbanized area, whereas a medium urbanized sample showed the greatest NoV GI concentration value of $1.69E+02$ GC/100L.

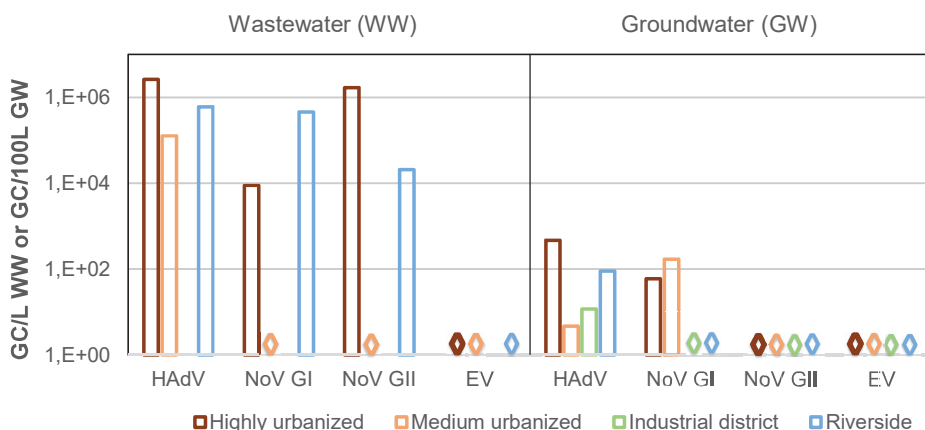


Figure 2. Viral pathogens detected and quantified in wastewater (WW) and groundwater (GW) samples from the winter campaign. Bars represent mean values of the concentrations obtained by qPCR assays. Samples that tested negative are indicated with $\Diamond$.

SARS-CoV-2 RNA was detected in winter wastewater samples using the N1 and N2 assays, with concentrations ranging between 2.39E+03 and 4.07E+05 CG/L. Groundwater, runoff, and summer wastewater samples all tested negative for SARS-CoV-2 assays.

The detection limits of the assays analyzing wastewater were 6.86E+03 GC/L for HAdV, 6.58E+03 GC/L for NoV GI, 4.74E+04 GC/L for NoV GII, 6.58E+03 GC/L for EV, and 2.95E+03 and 2.11E+03 for N1 and N2 SARS-CoV-2 targets, respectively. For groundwater samples, the detection limits were 1.71E+02 GC/100 L for HAdV, 1.65E+02 GC/100 L for NoV GI, 1.18E+03 GC/100 L for NoV GII, 1.65E+02 GC/100 L for EV, and 7.38E+01 and 5.28E+01 GC/100 L for N1 and N2 SARS-CoV-2 targets, respectively. For runoff samples, the detection limits were 2.14E+01 GC/L for HAdV, 2.06E+01 GC/L for NoV GI, 1.48E+02 for NoV GII, 2.06E+01 GC/L for EV, and 9.23E+00 and 6.60E+00 GC/L for N1 and N2 SARS-CoV-2 targets, respectively.

3.2 Virome of urban groundwater and runoff using TES

Urban groundwater and runoff samples were analyzed using the TES approach to explore the presence of vertebrate viruses in these matrices.

The viral assignments and reads sequenced in winter groundwater samples are presented in Figure 3, and additional details can be found in the Supplementary Material. Vertebrate viruses were sequenced in highly urbanized, industrial district, and riverside groundwater samples using TES. No viral reads were obtained from the medium urbanized groundwater sample; therefore, it was not included in the figure. Reads belonging to the Betapapillomavirus 2 species (Human papillomavirus, HPV-38, as the main identified type) were detected in the industrial district and riverside groundwater. Non-human vertebrate viruses were found in highly urbanized, industrial district, and riverside groundwater samples, with the highest diversity observed in the highly urbanized sample, including animal circoviruses and picornaviruses infecting bats, pigeons, or dogs. Bat circovirus was also detected in industrial district groundwater, while pigeon circovirus was identified in riverside groundwater.

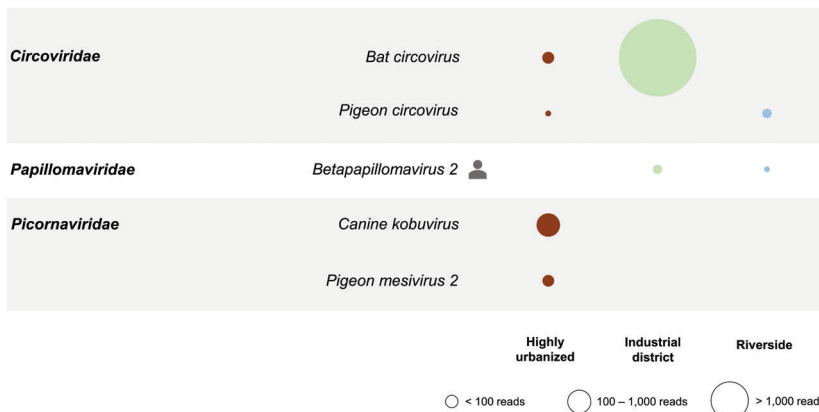


Figure 3. Vertebrate viruses and reads sequenced in highly urbanized, industrial district and riverside groundwater samples collected in winter sampling campaign. Human viruses are indicated with .

Figure 4 presents the viral assignments and reads sequenced in urban runoff samples, and additional details can be found in the Supplementary Material. A higher number of reads and number of viral assignments were obtained from RO1 sample, and most of the viral assignments in RO2 were found in RO1. A high diversity of human and other vertebrate viruses was identified in urban runoff, including members of *Anelloviridae*, *Astroviridae*, *Caliciviridae*, *Circoviridae*, *Coronaviridae*, *Genomoviridae*, *Parvoviridae*, *Picobirnaviridae*, *Picornaviridae*, *Poxviridae*, and *Retroviridae*. These assignments included viruses infecting pigeons, dogs, cats, or bats.



Figure 4. Vertebrate viruses and reads sequenced in medium urbanized runoff samples. Human viruses are indicated with .

4. DISCUSSION

The easy entry of pathogens into the urban environment through different pathways, such as the discharge of inadequately treated sewage, stormwater runoff, combined sewer overflows, and sanitary sewer overflows, is a significant concern for water resource managers (35). Groundwater is used for urban purposes, including irrigation for parks and gardens, street cleaning, etc. However, in recent years, cities have been enhancing the management of groundwater use with the objective of expanding its applications, considering the possibility of using it for human water supply.

In this study, qPCR assays revealed viral contamination in urban wastewater, runoff, and groundwater. The human fecal indicator HAdV was detected in 67% of groundwater samples, reaching concentrations of up to 4.67E+02 GC/100 L. The viral pathogen NoV GI was identified in winter samples from highly and medium urbanized groundwater, with values of 5.89E+01 and 1.69E+02 GC/100 L, respectively. In a recent review on waterborne viruses in urban groundwater environments, it was reported that over 80% of the studies in the literature identified NoV as the primary cause of the outbreaks. However, the occurrence of HAdV was higher in all studies where both were analyzed, suggesting that the evaluation of the presence of HAdV could be useful in predicting the presence of other waterborne pathogens in groundwater (1). The HAdV concentration in the runoff sample, 3.00E+00 GC/L, was consistent with concentrations reported in other studies and reviewed by Ahmed and colleagues (5).

There is limited literature available on groundwater virome. Using viral metagenomics, the viral families *Myoviridae*, *Siphoviridae* and *Podoviridae* have been described as the three most dominant families in groundwater (36,37), comprising viruses that infect bacteria and archaea. Although the applied approach in this study has been designed to capture sequences of vertebrate viruses, some bacteriophage sequences belonging to *Microviridae* were detected in riverside groundwater (Supplementary Material).

Waterborne viruses such as HAdV-41 and Hepatitis E Virus (HEV) have been detected in groundwater through untargeted viral metagenomics (13), and avian adenoviruses and sequences related to human and bovine alpha- and gamma- herpesvirus have also been found in sub-Saharan urban settlements (12). In addition, avian, equine and bat viruses belonging to *Genomoviridae* and *Parvoviridae*, as well as human viruses belonging to *Circoviridae* and *Parvoviridae* in groundwater used for irrigation in organic agriculture have been detected using the TES approach (30). In the current study, TES enabled the identification of human and other vertebrate viruses in urban groundwater. HPV-38, the main HPV type identified, has been linked to skin cancer (38). Groundwater samples also revealed other vertebrate viruses, including animal, circoviruses and picornaviruses infecting bats, pigeons, or dogs. Although the TES approach is known to facilitate the detection of vertebrate viruses, sequences from other viral hosts can be identified (13,30). Consequently, assignments of invertebrate and plant viruses, along with phages, were also obtained from the sequencing (Supplementary Material).

A higher diversity of viruses was identified in runoff samples using TES. Human and other vertebrate viruses were assigned as members of *Anelloviridae*, *Astroviridae*, *Caliciviridae*, *Circoviridae*, *Coronaviridae*, *Genomoviridae*, *Parvoviridae*, *Picobirnaviridae*, *Picornaviridae*, *Poxviridae*, and *Retroviridae* families. These assignments included viruses infecting pigeons, dogs, cats, or bats. As in the case of groundwater, invertebrate and plant viruses, and phages, were also identified in runoff samples (Supplementary Material). To the best of our knowledge, this is the first study employing NGS to explore the viral diversity in urban runoff samples. The presence of enteric viruses in runoff samples have been reported before by PCR or qPCR techniques, including adenoviruses, polyomaviruses, enteroviruses, and noroviruses (7,9). All these findings suggest that human sewage input could be the major source of fecal viral contamination in stormwater, posing a health risk due to potential exposure to a wide array of excreted pathogens.

Although groundwater samples tested positive for HAdV and NoV and a runoff sample for HAdV, no reads assigned to these viruses were obtained using TES. This could be attributed to the bias introduced by the SISPA step, performed before the sequencing library construction, that amplifies the most abundant genomes, underrepresenting others, as described before (30).

It is important to note that infectivity potential was not assessed in this study since qPCR and TES data does not provide information about the infectivity potential of the detected viruses. However, it is known that waterborne viruses can survive in groundwater environments for long periods and remain infective in water (39). Another significant limitation of this study is the limited number of sampling events. Future studies should focus on exploring viral diversity in these samples throughout the year to obtain a fully comprehensive overview of the groundwater and runoff virome, considering seasonal variability.

5. CONCLUSIONS

This study offers valuable insights into the virome present in urban groundwater and runoff. With increasing demands on water resources, the potential for surface and groundwater contamination by pathogens could result in an increase in waterborne disease outbreaks. The identification of a broad spectrum of pathogens that could contaminate groundwater, using targeted massive sequencing approaches, is crucial for assessing potential risks associated with waterborne diseases, especially in increasingly populated urban areas but also in rural areas that rely on groundwater use. Urban runoff, if contaminated with pathogenic microorganisms, may pose a significant risk. Implementing strategies to determine contamination sources and monitoring human and animal viruses with zoonotic potential in groundwater and runoff is essential to provide accurate information for assessing public health risks and develop effective mitigation measures.

DATA AVAILABILITY

The raw sequencing data generated during the current study are available in Zenodo under the DOI number 10.5281/zenodo.10514178.

ACKNOWLEDGMENTS

The authors would like to extend their gratitude to the staff of Barcelona Cicle de l'Aigua, SA (BCASA) involved in this study for their assistance and their technical support during the sampling campaigns.

REFERENCES

1. Rusiñol M. Waterborne viruses in urban groundwater environments. *PLOS Water*. 2023;2(8):e0000168.
2. Wallender EK, Ailes EC, Yoder JS, Roberts VA, Brunkard JM. Contributing Factors to Disease Outbreaks Associated with Untreated Groundwater. *Groundwater*. 2014;52(6):886–97.
3. Makepeace DK, Smith DW, Stanley SJ. Urban Stormwater Quality: Summary of Contaminant Data. *Crit Rev Environ Sci Technol*. 1995;25(2):93–139.
4. Graham KE, Anderson CE, Boehm AB. Viral pathogens in urban stormwater runoff: Occurrence and removal via vegetated biochar-amended biofilters. *Water Res*. 2021;207:117829.
5. Ahmed W, Hamilton K, Toze S, Cook S, Page D. A review on microbial contaminants in stormwater runoff and outfalls: Potential health risks and mitigation strategies. *Sci Total Environ*. 2019;692:1304–21.
6. Corsi SR, Borchardt MA, Spencer SK, Hughes PE, Baldwin AK. Human and bovine viruses in the Milwaukee River watershed: Hydrologically relevant representation and relations with environmental variables. *Sci Total Environ*. 2014;490:849–60.
7. Sidhu JPS, Ahmed W, Gernjak W, Aryal R, McCarthy D, Palmer A, et al. Sewage pollution in urban stormwater runoff as evident from the widespread presence of multiple microbial and chemical source tracking markers. *Sci Total Environ*. 2013;463–464:488–96.
8. Sidhu JPS, Hodggers L, Ahmed W, Chong MN, Toze S. Prevalence of human pathogens and indicators in stormwater runoff in Brisbane, Australia. *Water Res*. 2012;46(20):6652–60.

- 321 9. Sauer EP, VandeWalle JL, Bootsma MJ, McLellan SL. Detection of the human specific
322 Bacteroides genetic marker provides evidence of widespread sewage contamination of
323 stormwater in the urban environment. *Water Res.* 2011;45(14):4081–91.
- 324 10. Bernard K, Davis A, Simpson IM, Hale VL, Lee J, Winston RJ. Detection of SARS-CoV-2
325 in urban stormwater: An environmental reservoir and potential interface between human
326 and animal sources. *Sci Total Environ.* 2022;807.
- 327 11. Martínez-Puchol S, Rusiñol M, Fernández-Cassi X, Timoneda N, Itarte M, Andrés C, et al.
328 Characterisation of the sewage virome: comparison of NGS tools and occurrence of
329 significant pathogens. *Sci Total Environ.* 2020;713.
- 330 12. van de Vossenberg J, Hoiting Y, Kamara AK, Tetteh MK, Simaika JP, Lutterodt G, et al.
331 Double-stranded DNA virus assemblages in groundwater in three informal urban
332 settlements in sub-Saharan Africa differ from each other. *ACS Environ Sci Technol Water.*
333 2021;1(9):1992–2000.
- 334 13. Rusiñol M, Martínez-Puchol S, Timoneda N, Fernández-Cassi X, Pérez-Cataluña A,
335 Fernández-Bravo A, et al. Metagenomic analysis of viruses, bacteria and protozoa in
336 irrigation water. *Int J Hyg Environ Health.* 2020;224:113440.
- 337 14. Itarte M, Bofill-Mas S, Martínez-Puchol S, Torrell H, Ceretó A, Carrasco M, et al. Looking
338 for a needle in a haystack. SARS-CoV-2 variant characterization in sewage. *Curr Opin*
339 *Environ Sci Heal.* 2021;24:100308.
- 340 15. Departament d'Estadística i Difusió de Dades. Anuari estadístic de la ciutat de Barcelona
341 2020 [Internet]. Ajuntament de Barcelona. 2020 [cited 2024 Jan 16]. Available from:
342 <https://bcnroc.ajuntament.barcelona.cat/jsui/handle/11703/119961>
- 343 16. Forés E, Rusiñol M, Itarte M, Martínez-Puchol S, Calvo M, Bofill-Mas S. Evaluation of a
344 virus concentration method based on ultrafiltration and wet foam elution for studying
345 viruses from large-volume water samples. *Sci Total Environ.* 2022;829.
- 346 17. Forés E, Bofill-Mas S, Itarte M, Martínez-Puchol S, Hundesa A, Calvo M, et al. Evaluation
347 of two rapid ultrafiltration-based methods for SARS-CoV-2 concentration from wastewater.
348 *Sci Total Environ.* 2021;768.
- 349 18. Bofill-Mas S, Albinana-Gimenez N, Clemente-Casares P, Hundesa A, Rodriguez-
350 Manzano J, Allard A, et al. Quantification and stability of human adenoviruses and
351 polyomavirus JCPyV in wastewater matrices. *Appl Environ Microbiol.* 2006;72(12):7894–
352 6.
- 353 19. Hernroth BE, Conden-Hansson AC, Rehnstam-Holm AS, Girones R, Allard AK.
354 Environmental factors influencing human viral pathogens and their potential indicator
355 organisms in the blue mussel, *Mytilus edulis*: The first Scandinavian report. *Appl Environ*
356 *Microbiol.* 2002;68(9):4523–33.

- 357 20. Da Silva AK, Le Saux JC, Parnaudeau S, Pommepuy M, Elimelech M, Le Guyader FS.
358 Evaluation of removal of noroviruses during wastewater treatment, using real-time reverse
359 transcription-PCR: Different behaviors of genogroups I and II. *Appl Environ Microbiol.*
360 2007;73(24):7891–7.
- 361 21. Svraka S, Duizer E, Vennema H, De Bruin E, Van Der Veer B, Dorresteyn B, et al.
362 Etiological role of viruses in outbreaks of acute gastroenteritis in The Netherlands from
363 1994 through 2005. *J Clin Microbiol.* 2007;45(5):1389–94.
- 364 22. Hoehne M, Schreier E. Detection of norovirus genogroup I and II by multiplex real-time
365 RT-PCR using a 3'-minor groove binder-DNA probe. *BMC Infect Dis.* 2006;6:69.
- 366 23. Loisy F, Atmar RL, Guillon P, Le Cann P, Pommepuy M, Le Guyader FS. Real-time RT-
367 PCR for norovirus screening in shellfish. *J Virol Methods.* 2005;123:1–7.
- 368 24. Kageyama T, Kojima S, Shinohara M, Uchida K, Fukushi S, Hoshino FB, et al. Broadly
369 reactive and highly sensitive assay for Norwalk-like viruses based on real-time quantitative
370 reverse transcription-PCR. *J Clin Microbiol.* 2003;41(4):1548–57.
- 371 25. Rusiñol M, Hundesa A, Cárdenas-Youngs Y, Fernández-Bravo A, Pérez-Cataluña A,
372 Moreno-Mesonero L, et al. Microbiological contamination of conventional and reclaimed
373 irrigation water: Evaluation and management measures. *Sci Total Environ.* 2020;710.
- 374 26. CDC Division of Viral Diseases. 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-
375 PCR Diagnostic Panel:Instructions for Use. CDC-006-00019, Revis 08. 2023;
- 376 27. Briese T, Kapoor A, Mishra N, Jain K, Kumar A, Jabado OJ, et al. Virome capture
377 sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *MBio.*
378 2015;6(5).
- 379 28. Filipa-Silva A, Parreira R, Martínez-Puchol S, Bofill-Mas S, Barreto Crespo MT, Nunes M.
380 The Unexplored Virome of Two Atlantic Coast Fish: Contribution of Next-Generation
381 Sequencing to Fish Virology. *Foods.* 2020;9:1634.
- 382 29. Hjelmsø MH, Møllerup S, Jensen RH, Pietroni C, Lukjancenko O, Schultz AC, et al.
383 Metagenomic analysis of viruses in toilet waste from long distance flights—A new
384 procedure for global infectious disease surveillance. *PLoS One.* 2019;14(1).
- 385 30. Itarte M, Martínez-Puchol S, Forés E, Hundesa A, Timoneda N, Bofill-Mas S, et al. NGS
386 techniques reveal a high diversity of RNA viral pathogens and papillomaviruses in fresh
387 produce and irrigation water. *Foods.* 2021;10(8).
- 388 31. Martínez-Puchol S, Cardona L, Drago M, Gazo M, Bofill-Mas S. Viral metagenomics
389 reveals persistent as well as dietary acquired viruses in Antarctic fur seals. *Sci Rep.*
390 2022;12(1):1–11.

32. Mejías-Molina C, Pico-Tomás A, Beltran-Rubinat A, Martínez-Puchol S, Corominas L, Rusiñol M, et al. Effectiveness of passive sampling for the detection and genetic characterization of human viruses in wastewater. *Environ Sci Water Res Technol*. 2023;
33. Strubbia S, Schaeffer J, Besnard A, Wacrenier C, Le Mennec C, Garry P, et al. Metagenomic to evaluate norovirus genomic diversity in oysters: Impact on hexamer selection and targeted capture-based enrichment. *Int J Food Microbiol*. 2020;323:108588.
34. Kalantar KL, Carvalho T, De Bourcy CFA, Dimitrov B, Dingle G, Egger R, et al. IDseq-An open source cloud-based pipeline and analysis service for metagenomic pathogen detection and monitoring. *Gigascience*. 2020;9:1–14.
35. Arnone RD, Walling JP. Waterborne pathogens in urban watersheds. *J Water Health*. 2007;5(1):149–62.
36. Kallies R, Hölzer M, Toscan RB, da Rocha UN, Anders J, Marz M, et al. Evaluation of sequencing library preparation protocols for viral metagenomic analysis from pristine aquifer groundwaters. *Viruses*. 2019;11(6).
37. Zaouri N, Jumat MR, Cheema T, Hong PY. Metagenomics-based evaluation of groundwater microbial profiles in response to treated wastewater discharge. *Environ Res*. 2020;180:108835.
38. Kocjan BJ, Seme K, Cimerman M, Kovanda A, Potočnik M, Poljak M. Genomic diversity of human papillomavirus (HPV) Genotype 38. *J Med Virol*. 2009 Feb 23;81(2):288–95.
39. Abbaszadegan M, Lechevallier M, Gerba C. Occurrence of Viruses in US groundwaters. *J / Am Water Work Assoc*. 2003;95(9).

424

425 **AUTHOR CONTRIBUTION**

426 Conceptualization: MR, SB, EV, LS, MI. Methodology: MI, EF, SMP, MR, SB. Validation: MI, MR,
427 SB. Formal analysis: MI, EF, MR. Investigation: MI, EF, MR. Resources: MR, SB, EV, LS. Data
428 curation: MI, EF, SMP, MR, SB. Writing - original draft: MI, MR, SB. Writing – review and editing:
429 MI, MR, SB, SMP, EF, EV, LS. Visualization: MI, MR, SB. Supervision: MR, SB, EV, LS. Project
430 administration: MR, SB, EV, LS. Funding acquisition: MR, SB, EV, LS. All authors have read and
431 agreed to the published version of the manuscript.

432

433 **FUNDING**

434 This work was partially funded by the URBANWAT (PCI2019-103643) funded by the Spanish
435 Ministry of Science, Innovation and Universities and the UPWATER project (101081807) financed
436 by the EU call Horizon-CL6-2022-ZEROPOLLUTION-01. S. Bofill-Mas is a Serra-Hunter fellow at
437 the University of Barcelona. This study is performed with partial support from the Water Research
438 Institute of the University of Barcelona.

439

440 **COMPETING INTERESTS**

441 The authors declare no competing interests.

]

Capítol 3

Identificació de virus presents en
entorns de treball amb risc d'exposició.
Avaluació del risc ocupacional.



En aquest tercer i últim capítol de la tesi s'avalua l'exposició vírica en entorns laborals, ja que l'exposició ocupacional a patògens pot comportar riscos per a la salut dels treballadors. Aquest capítol consta d'un únic article, en el qual s'estudia l'exposició ambiental a virus en una EDAR i una granja de porcs mitjançant tècniques de NGS i estimacions de risc ocupacional mitjançant anàlisis de QMRA.

Article 6: Avaluació de l'exposició ambiental a virus en estacions depuradores d'aigua residual i granges de porcs mitjançant seqüenciació de nova generació i aproximacions de risc ocupacional

“Assessing Environmental Exposure to Viruses in Wastewater Treatment Plant and Swine Farm Scenarios with Next-generation Sequencing and Occupational Risk Approaches”

Itarte, M., Calvo, M., Martínez-Frago, L., Mejías-Molina, C., Martínez-Puchol, S., Girones, R., Medema, G., Bofill-Mas, S. i Rusiñol, M.

Journal of Hygiene and Environmental Health. Volume 259, June 2024, 114360.
<https://doi.org/10.1016/j.ijheh.2024.114360>

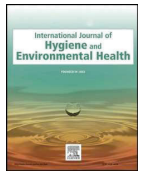
L'exposició ocupacional a patògens pot comportar riscos per a la salut dels treballadors. Aquest estudi investiga l'exposició vírica dels treballadors d'una EDAR i dels d'una granja de porcs mitjançant l'anàlisi de mostres d'aerosols i superfícies d'aquests llocs de treball. Les mostres d'aerosols es van obtenir amb el mostrejador d'aire ciclònic Coriolis μ , mentre que les superfícies es van mostrejar utilitzant adhesius de paper com a estratègia per mostrejar virus en superfícies a llarg termini. La contaminació vírica es va avaluar mitjançant assaigs de qPCR, i es va aplicar la tècnica de seqüenciació massiva amb enriquiment de dianes utilitzant VirCapSeq-VERT per identificar els virus de vertebrats als quals els treballadors podrien estar exposats. A més, es va realitzar una QMRA per estimar el risc ocupacional associat a l'exposició a virus per part dels treballadors de l'EDAR, seleccionant HAdV com a patògen de referència. A la granja de porcs, es va dur a terme una QMRA com a extrapolació, considerant un virus hipotèticament zoonòtic amb característiques similars a l'adenovirus porcí (PAdV). Les vies d'exposició dels models van

incloure la inhalació d'aerosols i la ingestió oral a través de superfícies contaminades i el contacte mà-boca.

HAdV i PAdV van ser detectats de manera extensa a les mostres de l'EDAR i de la granja de porcs, respectivament, mitjançant els assaigs de qPCR. La seqüenciació massiva utilitzant el panell de captura VirCapSeq-VERT va permetre va identificar virus humans i animals en mostres de l'EDAR, incloent-hi virus de les famílies *Adenoviridae*, *Circoviridae*, *Orthoherpesviridae*, *Papillomaviridae* i *Parvoviridae*. A la granja de porcs, la majoria dels virus de vertebrats identificats van ser virus de porc pertanyents a *Adenoviridae*, *Astroviridae*, *Circoviridae*, *Herpesviridae*, *Papillomaviridae*, *Parvoviridae*, *Picornaviridae* i *Retroviridae*.

Els resultats de la QMRA van revelar riscos considerables d'infeccions víriques per als treballadors de l'EDAR, si no es prenen mesures de seguretat. La probabilitat de malaltia per inhalació de HAdV va resultar ser més alta a l'estiu en comparació amb l'hivern, mentre que el major risc per ingestió oral es va observar als espais de treball durant l'hivern. La simulació de QMRA a la granja de porcs va suggerir un potencial risc considerable en cas d'exposició a un virus zoonòtic hipotètic.

Aquest estudi proporciona una visió d'interès sobre l'exposició ocupacional dels treballadors d'EDARs i granges de porcs a virus humans i animals. Els anàlisis de NGS i QMRA realitzats en aquest estudi poden ajudar els gestors a prendre decisions basades en evidències, facilitant la implementació de mesures de protecció i pràctiques per a la reducció de riscos per als treballadors.



Assessing environmental exposure to viruses in wastewater treatment plant and swine farm scenarios with next-generation sequencing and occupational risk approaches

Marta Itarte^{a,b,*}, Miquel Calvo^c, Lola Martínez-Frago^a, Cristina Mejías-Molina^{a,b}, Sandra Martínez-Puchol^a, Rosina Girones^{a,b}, Gertjan Medema^d, Sílvia Bofill-Mas^{a,b}, Marta Rusiñol^{a,b}

^a Laboratory of Viruses Contaminants of Water and Food, Secció de Microbiologia, Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona (UB), Barcelona, Spain

^b Institut de Recerca de l'Aigua (IdRA), Universitat de Barcelona (UB), Barcelona, Spain

^c Secció d'Estadística, Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona (UB), Barcelona, Spain

^d KWR Water Research Institute, Nieuwegein, Netherlands

ARTICLE INFO

Keywords:

Occupational exposure
Aerosol sampling
Sticker-based surface sampling
Target enrichment sequencing
Viral metagenomics
Quantitative microbial risk assessment

ABSTRACT

Occupational exposure to pathogens can pose health risks. This study investigates the viral exposure of workers in a wastewater treatment plant (WWTP) and a swine farm by analyzing aerosol and surfaces samples. Viral contamination was evaluated using quantitative polymerase chain reaction (qPCR) assays, and target enrichment sequencing (TES) was performed to identify the vertebrate viruses to which workers might be exposed. Additionally, Quantitative Microbial Risk Assessment (QMRA) was conducted to estimate the occupational risk associated with viral exposure for WWTP workers, choosing Human Adenovirus (HAdV) as the reference pathogen. In the swine farm, QMRA was performed as an extrapolation, considering a hypothetical zoonotic virus with characteristics similar to Porcine Adenovirus (PAdV). The modelled exposure routes included aerosol inhalation and oral ingestion through contaminated surfaces and hand-to-mouth contact.

HAdV and PAdV were widespread viruses in the WWTP and the swine farm, respectively, by qPCR assays. TES identified human and other vertebrate viruses WWTP samples, including viruses from families such as *Adenoviridae*, *Circoviridae*, *Orthoherpesviridae*, *Papillomaviridae*, and *Parvoviridae*. In the swine farm, most of the identified vertebrate viruses were porcine viruses belonging to *Adenoviridae*, *Astroviridae*, *Circoviridae*, *Herpesviridae*, *Papillomaviridae*, *Parvoviridae*, *Picornaviridae*, and *Retroviridae*.

QMRA analysis revealed noteworthy risks of viral infections for WWTP workers if safety measures are not taken. The probability of illness due to HAdV inhalation was higher in summer compared to winter, while the greatest risk from oral ingestion was observed in workspaces during winter. Swine farm QMRA simulation suggested a potential occupational risk in the case of exposure to a hypothetical zoonotic virus.

This study provides valuable insights into WWTP and swine farm worker's occupational exposure to human and other vertebrate viruses. QMRA and NGS analyses conducted in this study will assist managers in making evidence-based decisions, facilitating the implementation of protection measures, and risk mitigation practices for workers.

1. Introduction

People can be exposed to various risk factors in workplaces, with biological factors being of particular importance. Exposure to

bioaerosols in the occupational environment has garnered attention due to its association with various health effects, including infectious diseases, acute toxic effects, allergies, and cancer (Douwes et al., 2003). Bioaerosols are composed of small particles containing microorganisms

* Corresponding author. Laboratory of Viruses Contaminants of Water and Food, Secció de Microbiologia, Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona (UB), Barcelona, Spain.

E-mail address: mitarte@ub.edu (M. Itarte).

<https://doi.org/10.1016/j.ijheh.2024.114360>

Received 1 December 2023; Received in revised form 21 February 2024; Accepted 18 March 2024

1438-4639/© 2024 The Authors. Published by Elsevier GmbH. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

such as viruses, bacteria, and fungi, as well as organic compounds derived from microorganisms (Mandal and Brandt, 2011).

Numerous processes and activities can lead to the generation of bioaerosols in occupational settings, including manufacturing, agriculture, waste and wastewater treatment, farming, and laboratory environments (Mirskaya and Agranovski, 2018). In a Wastewater Treatment Plant (WWTP), most of the adverse health conditions reported among workers have been attributed to bioaerosols (Amoah et al., 2022; Carducci et al., 2016; Corrao et al., 2012). These aerosols can be generated by several wastewater treatment processes, and pathogenic microorganisms from wastewater can be easily released into the atmosphere in aerosols (Kataki et al., 2022).

Wastewater contains a diverse range of viruses, with geographical and seasonal differences observed among specific viral groups (Nieuwenhuijse et al., 2020). The abundance and diversity of pathogenic viruses in wastewater have been shown to reflect the pattern of infection in the human population, and some of the principal human pathogenic waterborne viruses include Human Adenovirus (HAdV), Rotavirus (RoV), Hepatitis A Virus (HAV), and other enteric viruses, such as Norovirus (NoV), Coxsackievirus (CV), and Astrovirus (AstV) (Corpus et al., 2020).

Aerosols of different sizes produced in WWTPs and all airborne biological agents can subsequently settle on surfaces (Han et al., 2013). Therefore, WWTP workers may be exposed to viruses through either the inhalation of bioaerosols generated during technological processes in the WWTP or oral ingestion after direct contact with contaminated surfaces, clothes, or tools. In fact, WWTP workers have been found to be at a higher risk of developing a wide range of work-related symptoms compared to the general population. These symptoms may include respiratory and gastrointestinal effects, such as diarrhea (Masclaux et al., 2014). The level of risk faced by WWTP workers also depends on factors such as the infectivity of the pathogen, its concentration, the duration of exposure, and the immunity (Carducci et al., 2016, 2018).

With the rapid expansion of large-scale and intensive swine production, the emission of aerosols from swine farms has also become a growing concern (Liu et al., 2023). This concern has intensified especially in the context of the One Health era, which emphasizes the interdependence of human, animal, plant, and environment health (One Health Commission, 2018). Aerosols produced in swine farms mainly originate from sources such as manure, feed, swine hair and skin, secondary production, and waste treatment (Liu et al., 2023). Airborne pathogenic viruses, including Influenza A Virus (IAV), Porcine Reproductive and Respiratory Syndrome Virus (PRRSV), Porcine Epidemic Diarrhea Virus (PEDV), Classical Swine Fever Virus (CSFV) and African Swine Fever Virus (ASFV), have been detected in the air of swine farms (Alonso et al., 2015; Corzo et al., 2013; Dee et al., 2009; Neira et al., 2016; O'Brien and Nonnenmann, 2016; Olesen et al., 2017; Weesendorp et al., 2009). Additionally, long-distance airborne transport of pathogens such as IAV and PRRSV has been demonstrated (Corzo et al., 2013; Dee et al., 2009) and IAV has also been detected on surfaces in swine production facilities during outbreaks (Neira et al., 2016). These findings imply potential risks, which align with data suggesting that swine workers and their non-swine-exposed families are at an increased risk of zoonotic influenza virus infections (Gray et al., 2007), leading to calls for the inclusion of these workers in pandemic surveillance and in antiviral and immunization strategies (Myers et al., 2006).

Next-generation sequencing (NGS) is a valuable and highly effective tool for detecting viral pathogens within a sample by simultaneously analyzing the sequences present. The metagenomic information obtained from these samples is essential for investigating potential biological exposures. To date, only a few studies have employed NGS techniques to characterize viral communities in aerosols and surfaces in the context of occupational exposure, such as aerosol samples from an animal slaughterhouse (Hall et al., 2013) and WWTP (Han et al., 2019), as well as surface and mobile phone swab samples from healthcare workers in the Emergency unit of a tertiary care facility (Boucherabine

et al., 2022). These studies utilized NGS to characterize the profile of viruses present in aerosol and surface samples, providing a deeper understanding of the viruses to which workers are exposed.

Quantitative Microbial Risk Assessment (QMRA) approaches can be used to evaluate the risk of infection within those occupational settings. The use of QMRA involves assessing the risk of harmful microorganisms through hazard identification, exposure assessment, dose-response modeling, and risk characterization (Haas et al., 1999). QMRA is commonly estimated through Monte Carlo simulations to quantitatively assess the range and probability of health risks (Chen et al., 2021; Lim et al., 2015; Shi et al., 2018; Yan et al., 2021). For risk characterization, the two primary health risk benchmarks widely used in describing the magnitude of QMRA outcomes are the acceptable annual infection risk level proposed by the U.S. Environmental Protection Agency (U.S. EPA, 2005) ($\leq 10E-4$ pppy, per-person-per-year) and the acceptable disability-adjusted life years (DALYs) proposed by the World Health Organization (WHO, 2008) ($\leq 10E-6$ DALYs pppy) (Chen et al., 2021; Yan et al., 2021). Additionally, the importance of all input variables can be identified through sensitivity analysis, which tests the relative impacts of stochastic input parameters on health risks (Chen et al., 2021). QMRA has been applied to assess the risk of viral infections associated with drinking water and food (Deere and Ryan, 2022; Petterson, 2016; Schijven et al., 2019), reclaimed water (Schoen et al., 2018), as well as in various working environments, such as WWTPs (Carducci et al., 2018; Dada and Gyawali, 2021; Medema et al., 2004; Zaneti et al., 2021), agricultural areas utilizing reclaimed wastewater for irrigation (Antwi-Agyei et al., 2016) or fertilizer applications (Brooks et al., 2005, 2012; Tanner et al., 2008), and scenarios dealing with leachate from municipal solid waste (MSW) (Lanzarini et al., 2022).

This study aims to provide valuable insights into the occupational exposure of workers to viruses in two different settings, WWTPs and swine farms. The first objective is to evaluate viral contamination through quantitative polymerase chain reaction (qPCR) assays and investigate the presence of vertebrate pathogenic viruses in bioaerosols and contaminated surfaces from these workplaces by employing NGS. To our knowledge, this is the first study to characterize the profiles of viruses infecting vertebrates present in aerosol and surface samples using a target enrichment sequencing (TES) approach. Additionally, the second objective of this study is to estimate the risk of exposure to viral pathogens with QMRA analysis in a WWTP, with HAdV as a model virus transmitted through mixed oral-fecal and respiratory routes. In the swine farm scenario, the risk is intended to be estimated through a QMRA simulation, considering the inhalation or oral ingestion of a hypothetical virus with assumed zoonotic potential and characteristics similar to Porcine Adenovirus (PAdV).

2. Material and methods

2.1. Sampling sites and samples

Two workplace scenarios located in Catalonia were chosen for this study: a WWTP and a swine farm. The selected WWTP is a conventional plant that treats wastewater from 183,517 inhabitants with a flow capacity of 57,000 m³/day. On the other hand, the selected swine farm is a breeding and gestation barn with an occupancy of 2,000 piglets, where newborn pigs spend the first weeks of life before being moved to a fattening farm. The sampling campaigns included both winter and a summer seasons, with three sampling events for each workplace scenario and season. Table 1 and Table 2 provide a summary of the exposure scenarios, samples analyzed, seasons, and the conducted analysis in the WWTP and the swine farm, respectively.

For aerosol sampling, a Coriolis μ air sampler (Bertin Technologies, France), designed to capture particles ranging from 0.5 to 20 μ m, was used at a flow rate of 300 L/min during 60 min in specific areas where a higher exposure to aerosols by workers was identified. Aerosol samples were collected in sterile collection cones, pre-filled with 15 ml of a saline

Table 1
WWTP exposure scenarios, specifying the type of samples analyzed, seasons and the type of analysis conducted.

		Samples analyzed	Season	Analysis
Workspaces	Centrifuge zone (indoor)	Aerosol	Winter	qPCR, NGS, QMRA
			Summer	qPCR, NGS, QMRA
		Surface	Winter	qPCR, NGS, QMRA
			Summer	qPCR, NGS, QMRA
	Reactor zone (outdoor)	Aerosol	Summer	qPCR, NGS
		Surface	Winter	qPCR, NGS, QMRA
	Other	Surface	Summer	qPCR, NGS, QMRA
			Winter	qPCR, NGS, QMRA
Break room		Surface	Summer	qPCR, NGS, QMRA
			Winter	qPCR, NGS, QMRA

Table 2
Swine farm exposure scenario, specifying the type of samples analyzed, seasons and the type of analysis conducted.

Samples analyzed	Season	Analysis
Aerosol	Winter	qPCR, NGS, QMRA
	Summer	qPCR, NGS, QMRA
Surface	Winter	qPCR, NGS, QMRA
	Summer	qPCR, NGS, QMRA

phosphate buffer solution (PBS). At the WWTP, aerosol samples were collected in the centrifuge zone (indoor) during winter (3 samples) and summer (4 samples). Additionally, in summer, 3 long-monitoring aerosol samples lasting 6 h were collected in the reactor zone (outdoor). At the swine farm, 3 aerosol samples were collected from the breeding area for each season. In all aerosol samplings, the manufacturer’s accessory to maintain a constant liquid volume and compensate evaporation was used.

For surface sampling, samples were collected following the strategy described by Sommer and coworkers as a long-term sampling approach to study viruses on surfaces (Sommer et al., 2021), technique previously outlined by Bobal and Witte for the detection and monitoring of bacterial pathogens (Bobal et al., 2019). Commercially paper-based stickers (Markierungspunkte ø 8 mm, permanent, ref. PSA08J 301; Avery of CCL Industries, Inc., Toronto, Canada), sterilized with UV-C radiation for a minimum of 15 min, were applied to frequently touched surfaces by workers, such as switches, doorknobs, and working tools, using sterile tweezers.

After 7 days, the stickers were removed and transferred to 2 ml Eppendorf tubes. At the WWTP, two scenarios were differentiated: the workspaces and the break room, which serves as the kitchen where workers take a break to eat. During winter, 12 surfaces samples from the workspaces (5 from the filtering zone, 2 from the centrifuge zone, 4 from the main building and 1 from the reactor zone) and 6 surface samples from the break room were collected on each sampling event of the season’s campaign. In summer, a selection of surface samples that showed higher contamination during winter was collected, consisting of 5 workspaces surface samples (2 from the filtering zone, 1 from the centrifuge zone, 1 from the main building and 1 from the reactor zone) and 2 break room samples. For the swine farm, each sampling event involved collecting 6 surface samples distributed on switches and doorknobs around the entire farm. Stickers were also collected from surfaces on upper walls, serving as a negative control for sampling and subsequent sample processing.

Both aerosol and surfaces samples were transported at 4 °C to the laboratory and processed on the same day of collection.

2.2. Sample processing and viral quantification

The concentration of viral particles from aerosol samples was performed by ultrafiltering the volume from the collection cones using the automatic Concentration Pipette (CP-Select™) with 150 kDa tips

(InnovaPrep, Drexel, MO, USA) to obtain a final volume of 300 µL, following the procedure previously described for viral concentration from wastewater (Forés et al., 2021). The viral nucleic acids from the concentrated aerosol samples were then extracted using the QIAamp® Viral RNA Mini Kit from QIAGEN (QIAGEN, Germantown, MD, USA). The extracted nucleic acids were eluted in 70 µL and stored at –80 °C for further analysis in qPCR and NGS assays.

For the surface samples, the viral nucleic acids were extracted using the RNeasy® PowerMicrobiome® Kit (QIAGEN, Germantown, MD, USA) with a final volume of 50 µL, following the manufacturer’s instructions. An additional pre-step of bead-beating for 30 s at 4 m s^{–1} using FastPrep-24™ (MP Bio, USA) was included, as previously described (Mejías-Molina et al., 2023). The extracted nucleic acids were stored at –80 °C for further analysis.

For evaluating the viral contamination in the WWTP and swine farm samples, specific qPCR assays were performed for Human Adenovirus (HAdV) and Porcine Adenovirus (PAdV), respectively. TaqMan® Environmental Master Mix 2.0 (Applied Biosystems, Waltham, MA, USA) was used, along with specific primers and probes previously described (Bofill-Mas et al., 2006; Hemroth et al., 2002; Hundersa et al., 2009). Additionally, all samples, from both the WWTP and the swine farm, were analyzed using the SARS-CoV-2 N1 assay (CDC Division of Viral Diseases, 2023) with the RNA UltraSense™ One-Step qRT-PCR System (Applied Biosystems, Waltham, MA, USA). With the exception of the synthetic SARS-CoV-2 control (control 51 from Twist Biosciences), the qPCR standards were prepared using synthetic gBlocks® Gene fragments (IDT, Coralville, IA, USA), which were quantified with a Qubit 3.0 dsDNA HS Assay Kit (Invitrogen, Carlsbad, CA, USA) and serially diluted from 10⁰ to 10⁷ copies per reaction. All qPCR assays were performed using the QuantStudio™ Real-Time PCR System from ThermoFisher Scientific. Both undiluted and 10-fold diluted nucleic acid extractions were analyzed, and non-template controls were included in each of the assays.

2.3. Metagenomic analysis: target enrichment sequencing (TES)

Nucleic acid extractions from aerosol and surface samples were pooled, considering the type of sample, the season of the year, and the workplace zone studied (further details can be found in the Supplementary Material, Table S1 and Table S2). The nucleic acid pools were analyzed using the NGS approach target enrichment sequencing (TES) to explore the diversity of vertebrate viral pathogens present in the aerosol and surface samples, aiming for a better characterization of the viruses to which WWTP and swine farm workers are exposed.

2.3.1. Sequence-independent, single-primer amplification (SISPA)

A total of 11 nucleic acid pools were prepared prior to library construction following a sequence-independent, single-primer amplification (SISPA) method, which has been previously described for studying both RNA and DNA viruses (Fernandez-Cassi et al., 2018, 2018b; Itarte et al., 2021; Martínez-Puchol et al., 2020; Mejías-Molina et al., 2023). Briefly,

a retrotranscription step was performed using a random nonamer primer and SuperScript IV (Invitrogen, Carlsbad, CA, USA), followed by a second-strand synthesis with Sequenase 2.0 (Applied Biosystems, Waltham, MA, USA). To obtain sufficient dsDNA for library construction, nucleic acids were amplified using AmpliTaq Gold DNA polymerase (Applied Biosystems, Waltham, MA, USA), and the resulting PCR products were purified with the Zymo DNA Clean & Concentrate kit (Zymo Research, Irvine, CA, USA). The purified PCR products were quantified using the Qubit 3.0 dsDNA HS Assay Kit (Invitrogen, Carlsbad, CA, USA).

2.3.2. Library construction

For each pool of samples, libraries were constructed using the KAPA HyperPlus Library Preparation Kit (KAPA Biosystems, Roche, Basel, Switzerland). Following the manufacturer's instructions, the library construction consisted of enzymatic fragmentation, end repair and A-tailing reaction, and adapter ligation. Then, samples were amplified using the KAPA UDI Primer Mixes (KAPA Biosystems, Roche, Basel, Switzerland). Afterwards, the resulting libraries were quantified using the Qubit 3.0 dsDNA HS Assay Kit (Invitrogen, Carlsbad, CA, USA).

2.3.3. Capture of viral sequences by VirCapSeq-VERT Capture Panel

Libraries were equimolarly pooled and captured using the VirCapSeq-VERT Capture Panel (Roche, Basel, Switzerland). This panel consists of probes designed to capture sequences from 207 viral taxa known to infect vertebrates through hybridization and a following post-capture PCR amplification, enabling the detection of viral sequences in complex sample types as described in previous studies (Briese et al., 2015; Filipa-Silva et al., 2020; Hjelmso et al., 2019; Itarte et al., 2021; Martínez-Puchol et al., 2020, 2022; Mejías-Molina et al., 2023; Strubbia et al., 2020). Captured libraries were sequenced using an Illumina NextSeq 500 platform for a 2 × 150 cycles Mid Output run.

2.3.4. TES bioinformatic processing

Bioinformatic analysis was performed using CZ ID portal (<https://czid.org/>, accessed on November 30, 2023), an open-source cloud-based pipeline and service for metagenomic pathogen detection and monitoring (Kalantar et al., 2020). Briefly, this portal performs host and quality filtration steps, followed by an assembly-based alignment pipeline, which results in the assignment of reads and contigs to taxonomic categories. Only viral assignments with a nucleotide identity above 70% and >100 bp alignment length were considered.

2.4. Quantitative Microbial Risk Assessment

The QMRA was constructed, as described in the following paragraphs, to estimate the individual-level occupational risk faced by workers in WWTPs and also as an extrapolation for workers in swine farms.

2.4.1. Risk assessment framework and hazard identification

The reference viruses HAdV and PAdV were selected for WWTP and

swine farm risk assessment, respectively. Both HAdV and PAdV are double-stranded DNA viruses belonging to the *Adenoviridae* family and they are commonly found in the environment due to their high resistance to environmental conditions (Hijnen et al., 2006). HAdV has been suggested as a human fecal indicator since it is persistently excreted in feces or urine by infected humans, both with and without clinical symptoms (Bofill-Mas et al., 2013), causing various conditions such as gastroenteritis, respiratory infections, eye infections, acute hemorrhagic cystitis, and meningoencephalitis (Allard and Vantarakis, 2017). On the other hand, PAdV has been used as a Microbial Source Tracking (MST) tool to identify sources of porcine fecal contamination (Rusiñol et al., 2014). Although PAdV infection in pigs is often subclinical, it can cause mild diarrhea, anorexia and dehydration (Kumthip et al., 2019).

PAdV is a virus that infects pigs and not humans but, as an assumption, it was chosen as a model of a hypothetical virus with zoonotic potential for the following reasons: (1) no potential zoonotic viruses were detected in these samples, (2) it was prevalently excreted and present on the farm, and (3) it allows the application of well-known parameters to construct a robust model (adenovirus parameters).

In the WWTP scenario, the QMRA analysis focuses on estimating the occupational risk of viral infections from pathogens present in wastewater. In contrast, the QMRA for the swine farm is an approximate analysis to assess the risk of workers getting infected with a hypothetical virus with an assumed zoonotic potential and characteristics similar to PAdV, in a One Health context. The exposure pathways considered for both scenarios include inhalation and oral ingestion, as these are known transmission routes for both HAdV and PAdV (Kumthip et al., 2019; Teunis et al., 2016).

2.4.2. Exposure assessment

The objective of the exposure assessment models was to estimate the dose of the selected virus to which WWTP and swine farm workers are exposed during their work-related tasks at different seasons of the year. The study considered two main exposure pathways: inhalation of aerosols and oral ingestion through contaminated surfaces. Additionally, for the WWTP scenario, a distinction was made between the surfaces in the workspaces and those in a break room where the workers take breaks to eat. This differentiation was made because it was considered that the workers' activities and behaviors in the break room were different compared to the workspaces.

2.4.2.1. Aerosol model. The daily HAdV or PAdV dose to which each WWTP or swine farm worker is exposed by inhalation (d_i) was estimated from Equation (1), adapted from (Carducci et al., 2018) and the input parameters detailed in Table 3.

$$d_i = C_{AdV} \times \frac{f_{conv}}{r_{eff}} \times t_{exp} \times r_{in} \times cf$$

(1)

In Equation (1), the concentration C_{AdV} represents the number of genomic copies (GC) of HAdV or PAdV per m³ detected in WWTP or swine farm air samples, respectively, and is expressed in GC/m³. These

Table 3
Input exposure parameters used in the QMRA model for the aerosol inhalation pathway.

Model inputs	Notation	Distribution/value	Unit	Source
Concentration of HAdV (WWTP) or PAdV (farm) in air	C_{AdV}	WWTP winter: Uniform (3.09, 6.34) WWTP summer: Uniform (86.13, 96.24) Farm winter: Uniform (1.53, 126.96) Farm summer: Uniform (6.15, 128.23)	GC/m ³	This study (see sections 2.1. and 2.2.)
Conversion factor	f_{conv}	1/700	TCID ₅₀ /GC	McBride et al. (2013)
Recovery efficiency	r_{eff}	WWTP: 80 Farm: 75	%	Estimated in section 2.4.3.2.
Exposure time	t_{exp}	WWTP: Uniform (60,120) Farm: Uniform (240,360)	minutes	Observed in this study (see section 2.4.3.3.)
Inhalation rate	r_{in}	Lognormal ($\mu = 1.4$, $\sigma = 0.51$)	m ³ /hour	EPA (2011)
Conversion factor	cf	1/60	hour/minutes	–

data were obtained in this study as explained before in sections 2.1. and 2.2. and follow a uniform distribution (see section 2.4.3.1. for boundary estimations). Since the qPCR data does not provide quantification of infectious particles, the constant $f_{conv} = 1/700$ is the ratio between tissue culture infective dose (TCID₅₀) and GCs, assuming 700 GC = 1 TCID₅₀ (McBride et al., 2013). The constant r_{eff} represents the recovery efficiency, which is the recoverable amount of HAdV or PAdV present in the sample. This efficiency slightly varies between the WWTP and the swine farm (see section 2.4.3.2. for estimations). The random variable t_{exp} corresponds to the amount of time that the worker is exposed to aerosols and is expressed in minutes per day, following a uniform distribution (see section 2.4.3.3.). The random variable inhalation rate r_{in} represents the inhalation rate, expressed in m³/hour, which is the volume of air inhaled per unit of time and it follows a lognormal distribution with $\mu = 1.4$ and $\sigma = 0.51$ (EPA, 2011). The constant conversion factor cf is equal to 1/60, which is used to transform minutes to hours.

2.4.2.2. Surface model. For the surface risk equations, considering oral ingestion through hand-to-mouth contact, the dose was estimated using the model from Lanzarini et al. (2022), a conceptual model developed by the Institute for Occupational Medicine (Cherrie et al., 2006; HSE et al., 2007), and adapted to the specific exposure scenarios of this study using the input parameters detailed in Table 4. Oral exposure of WWTP and swine farm workers relies on the oral ingestion of HAdV or PAdV, respectively, through the contact of contaminated hands with the mouth after touching contaminated surfaces with bare hands. Therefore, the daily exposure dose for the surface risk is given by:

$$d_s = C_{AdV} \times \frac{f_{conv}}{r_{eff}} \times f_{cont|surf} \times TE_{surf|hand} \times TE_{hand|mouth} \times f_{cont|mouth} \times n_{hours} \quad (2)$$

C_{AdV} represents the concentration of HAdV or PAdV expressed in GC/cm², data obtained in this study (sections 2.1. and 2.2.) and in this case, follows a lognormal distribution (see section 2.4.3.1. for the parameter estimations). The constant f_{conv} is also 1/700, and the constant r_{eff} is now 5.95%, experimentally determined (data not shown). The random variable $f_{cont|surf}$ represents the frequency of hand contact with any surface, and it follows a uniform distribution (see 2.4.3.3. for the estimation). The random variable $TE_{surf|hand}$ models the percentage of viral transfer efficiency from stainless steel surfaces to hands and follows a normal distribution (Lopez et al., 2013). $TE_{hand|mouth}$ is the constant percentage

of viral transfer efficiency from bare hands to mouth (Rusin et al., 2002). The number of hand-to-mouth contacts ($f_{cont|mouth}$) follows a discrete distribution (see section 2.4.3.3) and finally, n_{hours} represents the hours the worker spends in each studied workplace zone (workspaces or break room in WWTP and general workspaces on farm). The skin contact area between the hand and peri-oral compartment for each exposure event was assumed to be 1 cm².

2.4.3. Fitting parameters in aerosol and surface exposure models

2.4.3.1. Fitting the virus concentration parameters in aerosol and surfaces. As explained in the preceding subsection, both aerosol and surface concentrations in equations (1) and (2) are treated as random variables following a suitable probabilistic distribution. In the case of aerosol concentrations, where the number of samples was small, a uniform distribution was used (see Carducci et al., 2018) for both seasons and both scenarios. The parameters of these situations were trivially computed as the minimum and maximum values registered in each case. On the other hand, the number of surface samples was larger, enabling the adjustment of a lognormal distribution that fits the observed data. It is important to note that all the samples, including those recorded below the limit of detection, were used for these estimations. The limit of detection was determined based on the minimum value reported for each virus assay and each type of sample. Such samples were treated as censored data (see Canales et al. 2018), allowing the fitting of the lognormal distribution using the “fitdistrplus” package in R (R Core Team, 2023). The “fitdistr” method of this package fits a univariate distribution for censored data using maximum likelihood estimation.

2.4.3.2. Fitting the recovery efficiency in aerosol exposure model. A relationship between the diameter of the particles suspended in the air and the collection efficiency was established from the information provided in a white paper supplied by the air sampler’s manufacturer, Bertin Technologies (<https://www.bertin-technologies.com/>, accessed on November 30, 2023), about the particle collection efficiency according to size and flow rate. On the logarithmic scale, this relation appears to be approximately linear. Assuming that the diameter of the particles follows a lognormal distribution, the two parameters of the density can be obtained from Gholipour et al. (2021); Agranovski et al. (2004), which

Table 4
Input exposure parameters used in the QMRA model for the oral ingestion pathway from contaminated surfaces.

Model inputs	Notation	Distribution/value	Unit	Source
Concentration of HAdV (WWTP) or PAdV (farm) in surfaces	C_{AdV}	WWTP winter workspaces: Lognormal ($\mu = 1.42$, $\sigma = 3.16$) WWTP summer workspaces: Lognormal ($\mu = 0.21$, $\sigma = 3.61$) WWTP winter break room: Lognormal ($\mu = 0.89$, $\sigma = 3.13$) WWTP summer break room: Lognormal ($\mu = 0.51$, $\sigma = 2.31$) Farm winter: Lognormal ($\mu = 3.81$, $\sigma = 1.81$) Farm summer: Lognormal ($\mu = 9.17$, $\sigma = 2.57$)	GC/cm ²	This study (see sections 2.1. and 2.2.)
Conversion factor	f_{conv}	1/700	TCID ₅₀ /GC	McBride et al. (2013)
Recovery efficiency	r_{eff}	5.95	%	This study (data not down)
Frequency of hand-to-surface contacts	$f_{cont surf}$	WWTP workspaces: Uniform (6, 11) WWTP break room: Uniform (4.5, 10) Farm: Uniform (8, 12)	events/hour	Observed in this study (see section 2.4.3.3.)
Viral transfer efficiency from surface to bare hands	$TE_{surf hand}$	Normal (37, 16)	%	Lopez et al. (2013)
Viral transfer efficiency from hands to mouth	$TE_{hand mouth}$	34	%	Rusin et al. (2002)
Frequency of hand-to-mouth contacts	$f_{cont mouth}$	WWTP workspaces: $\mu \simeq 2.8$, $\sigma \simeq 2.7$ WWTP break room: $\mu \simeq 7.6$, $\sigma \simeq 3.6$ Farm: $\mu \simeq 2.8$, $\sigma \simeq 2.7$	events/hour	(see section 2.4.3.3.)
Number of hours spent at workplace	n_{hours}	WWTP workspaces: Uniform (2, 5) WWTP break room: Uniform (0.1, 1) Farm: Uniform (4, 6)	hour	Observed in this study (see 2.4.3.3.)

characterize the size particle range generated in WWTPs and swine farms, respectively. By applying the linear transformation, a mean recovery efficiency was obtained for the WWTP and the farm scenarios.

2.4.3.3. Fitting the exposure times and frequency of contacts in aerosol and surface exposure models. In the aerosol model, the parameters of the uniform distribution for t_{exp} were chosen based on the observed behavior of the workers. Similarly, in the surface model, the parameters for n_{hours} and $f_{cont|surf}$ were obtained in an analogous manner. Since these three variables were modelled as uniform distributions, with t_{exp} and n_{hours} being continuous, and $f_{cont|surf}$ being discrete, their boundaries were estimated as the minimum and maximum values in the observed data.

As for the frequency $f_{cont|mouth}$, the number of hand-to-mouth contacts was modelled as a discrete variable. The jump points and cumulative distribution were fitted to the quantiles in the adult's rows in the eating and non-eating macro-activity of Table 1 from Wilson et al. (2021).

2.4.4. Dose-response model

Dose-response models describe the relationship between exposure and the probability of infection and illness. The dose-response model developed by Teunis et al. (2016) was adopted in this study for HAdV and PAdV risk models, and it was used before in other QMRA studies (Carducci et al., 2018; Lanzarini et al., 2022). Briefly, this response model describes the distributions of infectivity and pathogenicity in adenovirus studies, incorporating differences in inoculation route (aerosol inhalation, oral ingestion and intranasal or intraocular droplet inoculation) as shift in average infectivity and pathogenicity. The daily dose variable allowed to compute the probability of infection (P_{inf}):

$$P_{inf}(d_s|\alpha, \beta) = 1 - {}_1F_1(\alpha, \alpha + \beta, -d_s) \quad (3)$$

where ${}_1F_1$ is the Kummer confluent hypergeometric function, and α and β are infectivity (infection dose-response) parameters of the distribution that refer to the inhalation route or the oral ingestion route. The infectivity parameters for the adenovirus dose-response were established as $\alpha = 5.24$ and $\beta = 2.95$ for the inhalation route (aerosol model), whereas $\alpha = 5.11$ and $\beta = 2.80$ were the parameters for the oral ingestion route (surface model) (Teunis et al., 2016). The daily probability of illness conditioned to be infected, denoted by $P_{ill|inf}$, was:

$$P_{ill|inf}(d_s) = 1 - \left(1 + \frac{d_s}{\eta}\right)^{-r} \quad (4)$$

where the pathogenicity parameters for the adenovirus dose-response used were $r = 3.04$ and $\eta = 3.36$ for the inhalation route (aerosol model), whereas $r = 0.41$ and $\eta = 6.53$ (Teunis et al., 2016) were the parameters for the oral ingestion route (surface model). Finally, the daily probability of illness (P_{ill}) was given by the following relation:

$$P_{ill}(d_s) = P_{ill|inf}(d_s) \times P_{inf}(d_s) \quad (5)$$

For any of the situations, the seasonal probability of illness was computed as:

$$P_{ill\ season} = 1 - \prod_{i=1}^{70} (1 - \text{Random}(P_{ill})) \quad (6)$$

where $\text{Random}(P_{ill})$ is a random sample from the distribution of P_{ill} (Karavarsamis and Hamilton, 2010). Both the winter and summer season were assumed to comprise 70 working days.

2.4.5. Risk characterization

Hazard identification, exposure assessment and dose-response assessment were combined to determine the probability of infection with HAdV in WWTPs and to provide an approximation for the probability of infection with a hypothetical virus with zoonotic potential and

characteristics similar to PAdV in swine farms. A Monte Carlo analysis was conducted for 10,000 simulations for each of the 10 different scenarios considered, obtaining their daily probability of illness and a suitable set of quantiles.

Following Lanzarini et al., 2022, a sensitivity analysis was conducted to assess the relative importance of the stochastic variables on the results of both models: the concentration C_{AdV} in both exposure models, the exposure time t_{exp} and the inhalation rate r_{in} in the aerosol model, whereas the transfer efficiency $TE_{surf|hand}$, the frequencies of contacts ($f_{cont|surf}$ hands-surface, and $f_{cont|mouth}$ hands-mouth) and the exposure time n_{hours} , in the surface model. To determine the effect of any variable -for instance, the variable v_i on the final risk estimate, all the other input parameters in the model equation were fixed to their average values, while v_i was treated as a random variable in the Monte Carlo runs. This computation was conducted over all the random variables v_i in the equation, fixing in each case all the parameters except the one of interest. Sensitivity analysis was also computed based on the median values, as some C_{AdV} distributions exhibit long-left-tailed asymmetric shapes.

All the simulations were conducted using R 4.3.0 (R Core Team, 2023).

3. Results

3.1. Viral quantification in aerosol and surface samples

Aerosol and surface samples from the WWTP and the swine farm were analyzed using specific qPCR assays for the detection and quantification of HAdV and PAdV, respectively. Additionally, all samples were analyzed for the presence of SARS-CoV-2. The qPCR assay results are summarized in Table 5 and Table 6, with further details available in the Supplementary Material (Table S1 and Table S2).

HAdV was detected in 60% of WWTP aerosol samples. In the winter season, all indoor samples from the centrifuge zone tested positive, with a mean concentration of $4.87 \text{ E}+00 \text{ GC}/\text{m}^3$. In the summer season, 2 out of 4 samples were positive with a mean value of $9.12 \text{ E}+01 \text{ GC}/\text{m}^3$, resulting in higher HAdV concentrations in summer centrifuge aerosol samples. For outdoor samples in the reactor zone during the summer, 1 out of 3 samples tested positive, with a concentration value of $9.04 \text{ E}+01 \text{ GC}/\text{m}^3$. Moreover, HAdV was also detected in 65% of workspaces surface samples and in 50% of the surface samples in the break room. Among these surface samples, those collected in the workspaces during winter had the highest HAdV concentrations, with a mean value of $1.29 \text{ E}+02 \text{ GC}/\text{cm}^2$.

PAdV was detected in 83% of the swine farm aerosol samples and 86% of the surface samples. Notably, all surface samples from summer tested positive and exhibited higher concentrations (mean value of $5.73 \text{ E}+04 \text{ GC}/\text{cm}^2$) compared to those reported in winter (mean value of $1.89 \text{ E}+02 \text{ GC}/\text{cm}^2$).

On the other hand, SARS-CoV-2 was not detected in any of the

Table 5

Quantification of HAdV in WWTP, respectively, during winter and summer seasons using qPCR, including mean concentrations (standard deviations, positive samples vs tested samples). HAdV concentrations in aerosol samples are expressed as GC/m^3 , while surface concentrations values are in GC/cm^2 . Additional details are available in the Supplementary Material (Table S1).

	Winter	Summer
Aerosol (GC/m^3)	Centrifuge zone (indoor): $4.87 \text{ E}+00$ ($1.64 \text{ E}+00$, 3/3)	Centrifuge zone (indoor): $9.12 \text{ E}+01$ ($7.15 \text{ E}+00$, 2/4) Reactor zone (outdoor): $9.04 \text{ E}+01$ (0, 1/3)
Surface (GC/cm^2)	Workspaces: $1.29 \text{ E}+02$ ($3.29 \text{ E}+02$, 23/32) Break room: $1.77 \text{ E}+01$ ($3.04 \text{ E}+01$, 8/18)	Workspaces: $2.94 \text{ E}+01$ ($2.38 \text{ E}+01$, 8/16) Break room: $1.27 \text{ E}+01$ ($1.35 \text{ E}+01$, 4/6)

Table 6

Quantification of PAdV in swine farm samples, respectively, during winter and summer seasons using qPCR, including mean concentrations (standard deviations, positive samples vs tested samples). PAdV concentrations in aerosol samples are expressed as GC/m³, while surface concentrations values are in GC/cm². Additional details are available in the Supplementary Material (Table S2).

	Winter	Summer
Aerosol (GC/m ³)	4.98 E+01 (6.75 E+01, 3/3)	6.72 E+01 (8.63 E+01, 2/3)
Surface (GC/cm ²)	1.89 E+02 (1.96 E+02, 14/19)	5.73 E+04 (7.42 E+04, 18/18)

aerosol and surface samples from the WWTP. In the case of the swine farm, two surface samples tested positive for the presence of SARS-CoV-2 RNA, which was linked to an outbreak among the workers (further discussed in the discussion section). The viral concentrations quantified in these samples were 1.35 E+01 and 2.32 E+02 GC/cm².

3.2. NGS results

Aerosol and surface samples from the WWTP and the swine farm were analyzed using the TES approach to explore the presence of vertebrate viruses within these environments. For this purpose, nucleic acids extractions were pooled, considering the type of sample, season of the year, and workplace zone studied. The TES approach involved capturing sequences from vertebrate viruses during library preparation by employing the VirCapSeq-VERT Capture Panel (Roche, Basel, Switzerland).

The viral assignments and reads sequenced in WWTP samples are presented in Fig. 1, and additional details can be found in the Supplementary Material (Tables S3–S9). TES enabled the detection of human and other vertebrate-infecting viruses in aerosol and surface samples from the WWTP. The sequencing reads obtained from the WWTP were assigned to viruses from *Adenoviridae*, *Circoviridae*, *Orthoherpesviridae*, *Papillomaviridae*, and *Parvoviridae* families. Additionally, several families, including *Anelloviridae*, *Astroviridae* (Human astrovirus), *Caliciviridae* (Norwalk virus), *Coronaviridae* (Avian coronavirus) and *Picornaviridae*, were only sequenced in aerosol samples, while *Herpesviridae* (Porcine lymphotropic herpesvirus 3), *Polyomaviridae* (Merkel cell polyomavirus, MCPyV) and *Retroviridae* were exclusively found in surface samples. The highest viral diversity, considering the number of different viral assignments, was observed at the indoor samples (centrifuge zone) collected in the summer months, followed by workspaces surfaces in winter.

The *Parvoviridae* family accounted for the highest number of reads at the indoor aerosol samples in both seasons, as well as at the surfaces of workspaces in winter with Adeno-associated virus accounting for the majority of reads. Indoor aerosols from summer presented a high viral diversity within this family including a variety of known hosts susceptible to these viruses, such as humans, birds, dogs, pigs, and rodents.

The highest diversity of *Circoviridae* was also found in summer indoor aerosols, including human circovirus and cyclovirus, and pigeon circovirus.

A notable viral diversity of *Papillomaviridae*, *Orthoherpesviridae* and *Retroviridae* members was detected in surface samples, in both workplaces and the break room. Most of the *Papillomaviridae* reads were assigned to alpha, beta, gamma and mu Human Papillomavirus (HPV) genera, and also to Canine Papillomavirus, Epstein-Barr virus (EBV), belonging to *Orthoherpesviridae* family, was the most ubiquitous viral assignment within this viral family, as it was also detected in most of the surface samples. *Retroviridae*-assigned reads belonged to avian, sheep and goat viruses. MCPyV reads were detected only in workspace surfaces in winter.

HAdV, used as a reference virus in the QMRA analysis, was detected in all aerosol and surface samples, except for the break room samples, and the main serotype identified in all samples was HAdV-41. Other members of *Adenoviridae* family were detected, such as Fowl adenovirus

in summer aerosol and workspaces surface samples and Turkey adenovirus only in the break room surfaces in winter.

Fig. 2 presents the viral assignments and reads obtained from swine farm samples, and additional information is available in the Supplementary Material (Tables S10–S13).

TES also enabled the detection of vertebrate viruses in aerosol and surface samples from the swine farm. Most of the sequenced reads in aerosol and surface samples were assigned to porcine viruses, belonging to families like *Adenoviridae* (PAdV), *Astroviridae*, *Circoviridae*, *Herpesviridae*, *Papillomaviridae*, *Parvoviridae*, *Picornaviridae* and *Retroviridae* (Porcine type-C oncovirus). Additional families, including *Anelloviridae* (Torque teno sus virus), *Genomoviridae* (Porcine feces-associated gemycircularvirus) and *Tobamiviridae* (Bovine torovirus and Porcine torovirus), were exclusively detected in aerosol samples, while *Polyomaviridae* and *Sedoreoviridae* members (Porcine polyomavirus and Porcine Rotavirus A, respectively) were only found in surface samples. The greatest viral diversity was observed in winter aerosol samples, followed by summer surface samples.

Porcine bocavirus (PBoV) was the viral assignment that accounted for the highest number of reads in aerosol samples, whereas in surface samples, it was Porcine type-C oncovirus. Aerosol samples from winter and summer showed a high similarity in viral assignments between them. The aerosol sample collected in winter only showed few additional viral assignments compared to the summer sample, including PAdV, Porcine circovirus, Porcine feces-associated gemycircularvirus, Adeno-associated virus, Porcine parvovirus, and Porcine enterovirus. Similarly, high similarity was observed between surface samples from winter and summer, with the summer sample having a few more viral assignments, including Porcine astrovirus, Porcine stool-associated circular virus, Gamapapillomavirus, Porcine parvovirus, the bat-infecting virus *Rhinolophus sinicus* bocaparcovirus 2, Porcine enterovirus and Porcine Rotavirus A. In contrast, Porcine circovirus was the only viral assignment detected in winter samples.

The reference virus PAdV, used in the swine farm QMRA simulation as a model virus, was detected in all surface samples and in the winter aerosol sample.

3.3. QMRA

3.3.1. QMRA analysis in the WWTP

Fig. 3 presents the distributions of the daily illness probability (P_{ill}) for a WWTP worker, considering the inhalation or oral ingestion of HAdV. The scenario involving inhalation of HAdV through aerosols in the reactor zone (outdoor) was not included in the QMRA analysis because only one sample tested positive for HAdV, which did not provide sufficient data for the model.

When inhaling aerosols in the centrifuge zone (indoor), the highest probability of illness occurs during the summer season, with a mean value of 5.4%. In contrast, the highest probability of illness resulting from oral ingestion due to contaminated surfaces and hand-to-mouth contact was observed in the workspaces during the winter, with mean and median values of 13% and 0.14%, respectively. Surfaces in the workspaces exhibited a higher probability of illness compared to those in the break room.

The mean seasonal probability of illness for a WWTP worker exceeded 90% for both HAdV transmission routes in both seasons, expect for oral ingestion from contaminated surfaces in the break room and inhalation during winter (88.48% and 1.47%, respectively).

3.3.2. QMRA simulation in the swine farm

The distributions of daily illness probability (P_{ill}) for a swine farm worker, considering the inhalation or oral ingestion of a hypothetical virus with an assumed zoonotic potential and with characteristics similar to PAdV, are depicted in Fig. 4. The graph suggests that, assuming the presence of a potential zoonotic virus with similar characteristics to PAdV, the highest probability of illness would occur



Fig. 1. Viral assignments and reads sequenced in aerosol and surface samples from the WWTP using target enrichment sequencing. Centrifuge zone aerosols (indoor) were analyzed in both winter and summer, while reactor zone aerosols (outdoor) were only analyzed in summer. Regarding surface samples, which were also analyzed in both winter and summer, two scenarios were distinguished: the workspaces and the break room.

through oral ingestion from contaminated surfaces, particularly during the summer season, with mean and median risk values of 63.6% and 87.5%, respectively. The probability of illness due to the hypothetical virus inhalation would be similar between winter and summer, with mean values of 20.8% and 21.8%, respectively, both of which are lower

than the values associated with the oral ingestion route. As an approximation, the mean seasonal probability of illness for a swine worker, in the hypothetical scenario of the presence of a virus with zoonotic potential and characteristics similar to PAdV, would be 100% for both transmission routes in both seasons.

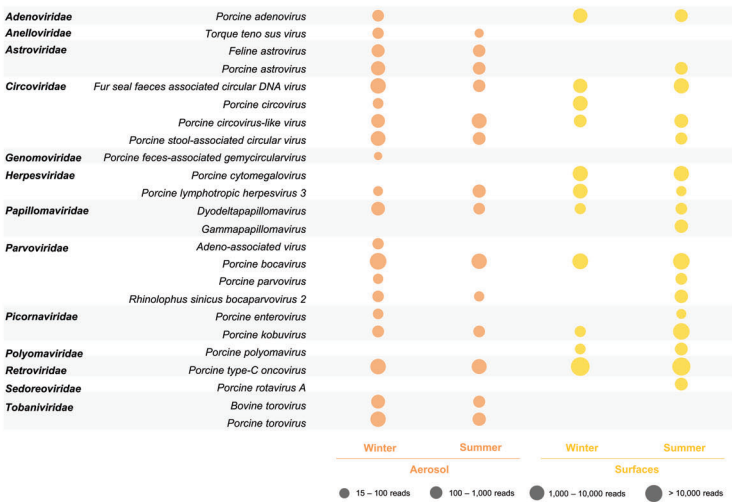


Fig. 2. Viral assignments and reads sequenced in aerosol and surface samples from the swine farm using target enrichment sequencing. Both aerosol and farm samples were analyzed in winter and summer.

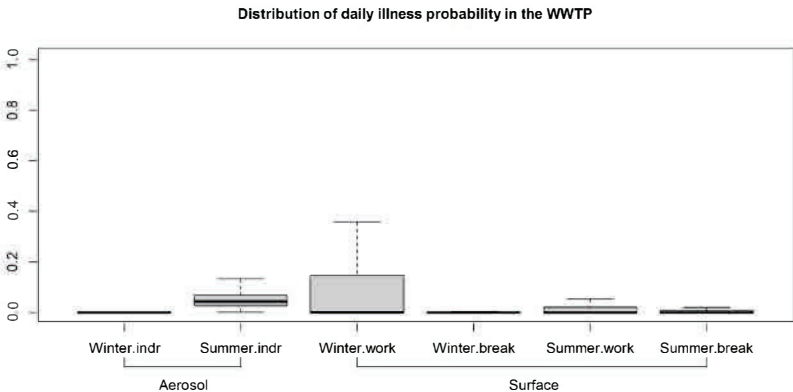


Fig. 3. Boxplots showing the distribution of daily illness probability in the WWTP, considering inhalation of HAdV from indoor aerosols in winter (Winter.indr) and summer season (Summer.indr), as well as oral ingestion of HAdV from contaminated surfaces in the workspaces and the break room in winter (Winter.work, Winter.break) and summer seasons (Summer.work, Summer.break).

3.3.3. Sensitivity analysis

Table 7 provides a summary of sensitivity analysis results, in terms of variance proportion of P_{ill} for the aerosol and surface models across the different scenarios and seasons. In the aerosol model, the analysis revealed that the parameter with the greatest influence on the model output's P_{ill} variations was the inhalation rate (r_{in}) in WWTP scenarios. On the other hand, in swine farm, P_{ill} was most sensitive to variations in the concentration of PAdV (C_{PAdV}) parameter.

In the surface model, the parameter with the highest influence was the number of hand-to-mouth contacts ($f_{cont(mouth)}$) in nearly all scenarios for both WWTP and the swine farm, except for the break room in the WWTP during both seasons, where the concentration of HAdV (C_{HAdV}) parameter held the greatest influence.

4. Discussion

Using suitable sampling and monitoring tools is essential for virus detection and understanding their transmission routes. The Coriolis μ air

sampler enabled the detection of viruses in both indoor and outdoor environments at the WWTP, as well as in the swine farm. Other studies have also employed the Coriolis μ sampler to detect viruses in WWTP (Brisebois et al., 2018; Stobnicka-Kupiec et al., 2022) and in swine, cattle and poultry farms (Prost et al., 2019; Salem et al., 2019; Scoizec et al., 2018).

Surface samples were collected following the long-term strategy described by Sommer and colleagues (Sommer et al., 2021). In the present study, this paper-based sticker approach demonstrated to be a successful method for detecting viruses on surfaces in the WWTP, as well as in the swine farm, using qPCR and NGS assays. A few prior studies have examined the presence of viruses on WWTP and swine farm surfaces, with most utilizing swabs for sampling (Anderson et al., 2021; López-Lorenzo et al., 2022; Prost et al., 2019; Stobnicka-Kupiec et al., 2022) while Neira and colleagues employed a sterile gauze dipped in their sampling procedures (Neira et al., 2016). To the best of our knowledge, this is the first study to utilize paper-based sticker strategy for surface sampling and subsequently perform NGS and QMRA analysis.

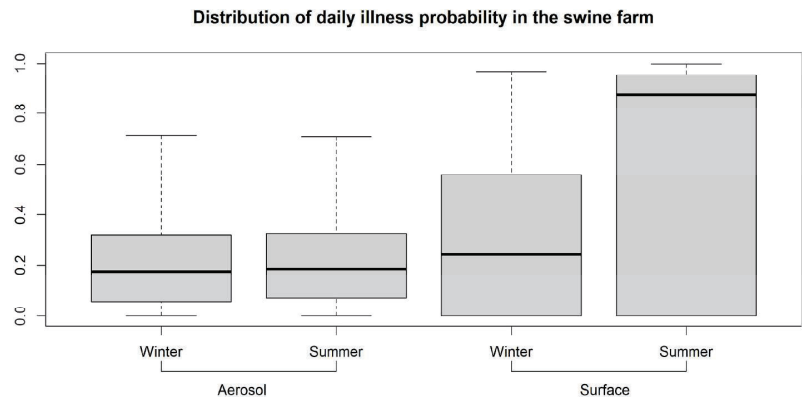


Fig. 4. Boxplots showing the distribution of daily illness probability in the swine farm, considering inhalation or oral ingestion from contaminated surfaces of a hypothetical virus with an assumed zoonotic potential and characteristics similar to PAdV.

Table 7
Sensitivity analysis in terms of variance proportion for each input parameter in the aerosol and surface models that affects the change of the probability of illness. Aerosol model (a) parameters: C_{Adv} (concentration of HAdV, WWTP, or PAdV, farm, in air), t_{exp} (exposure time) and r_{in} (inhalation rate); scenarios: WWTP.wtr.indr (WWTP winter indoor), WWTP.smr.indr (WWTP summer indoor), Farm.wtr (farm winter) and Farm.smr (farm summer). Surface model (b) parameters: C_{Adv} (concentration of HAdV, WWTP, or PAdV, farm, in surface), $TE_{surf|hand}$ (viral transfer efficiency from surface to bare hands), $f_{cont|surf}$ (frequency of hand-to-surface contacts), $f_{cont|mouth}$ (frequency of hand-to-mouth contacts) and n_{hours} (number of hours spent at workplace); scenarios: WWTP.wtr.wrk (WWTP winter work spaces), WWTP.wtr.break (WWTP winter break room), WWTP.smr.wrk (WWTP summer work spaces), WWTP.smr.break (WWTP summer break room), Farm.wtr (Farm winter) and Farm.smr (Farm summer).

a) Aerosol						
	WWTP.wtr.indr		WWTP.smr.indr			
			Farm.wtr	Farm.smr		
C_{Adv}	10.34		70.90	70.77		
t_{exp}	9.62		3.58	4.05		
r_{in}	49.67		32.08	34.32		
b) Surface						
	WWTP.wtr.wrk	WWTP.wtr.break	WWTP.smr.wrk	WWTP.smr.break	Farm.wtr	Farm.smr
C_{Adv}	116.46	105.9	123.32	105.05	70.76	10.26
$TE_{surf hand}$	18.81	47.1	29.38	27.82	12.87	1.28
$f_{cont surf}$	0.62	11.8	0.99	8.02	0.23	0.0005
$f_{cont mouth}$	204.90	58.9	296.56	36.30	115.42	112.47
n_{hours}	1.43	60.3	2.30	35.65	0.23	0.0004

All in all, both aerosol and surface sampling strategies allowed the detection and quantification of viral contamination using qPCR assays in samples from the WWTP and the swine farm, and the detection of sequences from vertebrate-infecting viruses through NGS analysis using TES. The viral contamination of aerosol and surface samples from the WWTP and the swine farm was evaluated by analyzing the prevalence of HAdV and PAdV, respectively. HAdV was detected in both aerosol and surface samples from the WWTP. This human virus is known to be shed in high concentrations and exhibits stability under various environmental conditions and disinfection treatments (Allard and Vantarakis, 2017). HAdV was detected in 60% of the WWTP aerosol samples. Samples collected from the centrifuge zone (indoor) exhibited higher HAdV concentration during summer (mean value of 9.12×10^1 GC/ m^3) compared to winter, which is consistent with previous findings, (Masclaux et al., 2014) and it has also been described for other microbial aerosols in WWTP (Grisoli et al., 2009; Oppliger et al., 2005). This could be related to the expected high evaporation of the wastewater and the possible increased emission of bioaerosols. HAdV was also detected in aerosol samples from the reactor zone (outdoor) at lower concentrations, as expected, due to its location in an open-air area. Other studies have detected higher concentration of HAdV in indoor WWTP air samples and observed slower concentrations outdoors (Carducci et al., 2016; Masclaux et al., 2014).

As for surfaces, HAdV was detected in 65% of the WWTP workspaces

samples and in 50% of the surface samples in the break room. In contrast to aerosol samples, the surface samples collected from the WWTP workspaces in summer exhibited lower concentrations of HAdV compared to winter. One possible explanation for this is that since most of the workspace surface samples were located outdoors and exposed to direct sunlight for an extended period, factors such as intensive solar radiation and high temperatures could have reduced the viral load on the paper-based stickers.

A recent investigation has explored the use of qPCR in combination with propidium monoazide (PMA) dye pretreatment, also known as viability-PCR (v-qPCR), as a method to distinguish potentially infectious and non-infectious viral particles at WWTP workplaces (Stobnick-a-Kupiec et al., 2022). The study reported a mean concentration of potentially infectious AdV of 4.41×10^3 GC/ m^3 in air samples, as well as 2.22×10^1 GC/ cm^2 on steel and 6.68×10^0 GC/ cm^2 on plastic swab surfaces samples. These findings suggest that HAdV detected in WWTP aerosol and surface samples in this study could be potentially infectious, although infectivity potential was not assessed.

In the swine farm, PAdV was detected in both aerosol and surface samples. PAdV is commonly found within swine populations, present in feces, residual water and sludge (De Motes et al., 2004; Hundesa et al., 2006; Rusiñol et al., 2014). PAdV is primarily transmitted through inhalation and the fecal-oral route, causing asymptomatic or mild illness in pigs. Typical symptoms include watery diarrhea, dehydration,

depression, and vomiting. In some cases, PAdV infection can lead to glomerulonephritis, chronic pneumonitis, and even death (Benfield and Richard, 2012; Kumthip et al., 2019; Nietfeld and Leslie-Stein, 1993). PAdV was detected in 83% of the swine farm aerosol samples, with mean concentrations values of $4.98 \text{ E}+01 \text{ GC/m}^3$ in winter and $6.72 \text{ E}+01 \text{ GC/m}^3$ in summer. PAdV detection in aerosol samples from a pig handling facility has been described before using conventional PCR (Poh et al., 2017) but to our knowledge, no data on PAdV concentration in swine farm aerosols have been reported to date. Regarding surfaces, PAdV was detected in 86% of the surface samples. All surface samples from summer tested positive and exhibited higher concentrations of PAdV (mean value of $5.73 \text{ E}+04 \text{ GC/cm}^2$) compared to those reported in winter (mean value of $1.89 \text{ E}+02 \text{ GC/cm}^2$). The presence of PAdV genetic material on swine farm surfaces may result from the deposition of airborne PAdV, as well as contact from contaminated hands. Upon observing the behavior of workers, it was noted that workers tended to wear gloves during winter but not during summer. This could provide a possible explanation for the higher PAdV concentrations detected on surfaces in the summer season. To date, studies analyzing the presence of viruses on surfaces in swine farms have primarily focused on Influenza A virus, porcine circovirus type 2 (PCV2), and porcine circovirus type 3 (PCV3) (Anderson et al., 2021; López-Lorenzo et al., 2022; Neira et al., 2016; Prost et al., 2019). Like aerosols, the concentration of PAdV on swine farm surfaces has not been previously described.

SARS-CoV-2 RNA was detected in two surface samples collected from the swine farm. Importantly, both samples were obtained on a day when there was an ongoing outbreak of the virus among the workers. While it appears that the risk of SARS-CoV-2 transmission through contaminated surfaces is low, the role of fomite transmission has been widely discussed, and the plausibility of a SARS-CoV-2 transmission chain from fomites has been suggested (Onakpoya et al., 2022). Furthermore, the susceptibility of pigs to SARS-CoV-2 is a subject of controversy (Frazzini et al., 2022). In this study, NGS data were obtained using a targeted assay (TES) with the VirCapSeq-VERT Capture Panel (Roche, Basel, Switzerland) to characterize the presence of vertebrate viral pathogens. This panel comprises approximately 2 million biotinylated oligonucleotide probes designed to bind coding sequences of all known viral taxa that infect vertebrates. While TES has been successfully employed in previous studies (Briese et al., 2015; Filipa-Silva et al., 2020; Hjelmsø et al., 2019; Itarte et al., 2021; Martínez-Puchol et al., 2020, 2022; Mejías-Molina et al., 2023; Strubbia et al., 2020), this is the first study to utilize this approach for the analysis of aerosol and surface samples.

The TES approach allowed the detection of vertebrate viruses in aerosol and surface samples from both the WWTP and the swine farm. In the case of the WWTP, the reference pathogen HAdV was detected in all aerosol and surface samples from the WWTP, except for the break room samples. The primary serotype detected was HAdV-41. Interestingly, HAdV-41 was also the predominant serotype identified through HAdV nested PCR and subsequent Sanger analysis of wastewater samples collected during the same period as the aerosol and surface sampling events in the WWTP (data not shown). HAdV-41 is etiologically associated with gastroenteritis and its high prevalence in environmental samples has been identified previously (Bofill-Mas et al., 2013). In addition to HAdV, other human and vertebrate viruses belonging to different viral families were detected in aerosol and surface samples from the WWTP. In the study conducted by Han et al. (2019), the metagenomic analysis of viral population in submicron aerosols emitted during wastewater treatment resulted in the detection of viruses primarily having bacteria or archaea as natural hosts. In this study, the TES approach facilitated the detection of vertebrate-infecting viruses, including HAdV, Human astrovirus, Norwalk virus, Human circovirus, Human cyclovirus, beta-HPV, Adeno-associated virus, and Human bocavirus. These viruses have been described as excreted and waterborne viruses, reinforcing that the source of these viruses would be the aerosolization from the wastewater. Some of these viruses cause asymptomatic infections and also outbreaks or sporadic cases with a

wide range of symptoms, from mild to severe gastroenteritis to meningitis, respiratory disease, conjunctivitis, myocarditis, paralysis, or hepatitis (Rusiñol and Girones, 2017). HPVs were initially believed to be mainly epitheliotropic, but some studies have reported their presence in raw sewage and river waters, including high- and low-risk oncogenic types (Iaconelli et al., 2015; Itarte et al., 2021; La Rosa et al., 2013; Martínez-Puchol et al., 2020; Rusiñol et al., 2020). Additionally, evidence of their excretion in the feces of patients with diarrhea has been documented (Di Bonito et al., 2015; Iaconelli et al., 2015; La Rosa et al., 2013), suggesting a potential transmission route through fecal shedding. A diverse array of avian viruses was detected in both aerosol and surface samples, aligning with prior studies that found avian viruses in wastewater samples from the same WWTP and confirming the impact of poultry industry effluents (Carratalà et al., 2012).

If the viruses detected in aerosol samples are suspected to originate from wastewater aerosolization, the profile of viruses found in surface samples provided insights suggesting that the presence of most of these viruses may be more likely caused by the contaminated hands of WWTP workers. A variety of different genera of HPVs were detected in all surface samples from the WWTP. HPVs infect the skin and mucosa epithelia, with effects ranging from benign lesions, such as common warts, to malignant carcinomas (Egawa et al., 2015). In most of the surface samples, EBV was also detected. This virus, mainly transmitted through saliva, is highly prevalent in the human population and can cause latent infection, being associated with many diseases, from mild asymptomatic infection to tumorigenesis (Huang et al., 2023). Finally, another human virus detected in workspaces surfaces in winter was MCPyV. This virus is the first polyomavirus to be associated with human cancer (Liu et al., 2018). It is highly prevalent in the general population, and nearly all healthy adults asymptotically shed MCPyV from their skin. However, this infection can lead to a lethal form of skin cancer in elderly and immunosuppressed individuals (Liu et al., 2016). Further investigation should be conducted to explore the seasonality of these viruses. The SISPA protocol, performed before library preparation to overcome the limitation of low quantities of viral genomes, may introduce bias by amplifying sequences in a random manner, as previously reported (Duhaime et al., 2012; Itarte et al., 2021; Karlsson et al., 2013). Therefore, additional research is required to elucidate whether the variations in virus detection between seasons are attributable to SISPA bias or are due to the seasonal behavior of the viruses.

Regarding the swine farm, the TES approach also enabled the detection of PAdV, in all surface samples and in the winter aerosol sample. PAdV infections are common in swine, and this virus appears to persist in swine tissues for extended periods (Paul et al., 2003). Additionally, other viruses were detected in both aerosol and surface samples from the swine farm, with most of them being described as common viruses infecting swine and belonging to different viral families. Among the detected viruses, PBoV was the viral assignment with the highest number of reads in aerosol samples. PBoV has been reported worldwide, primarily in weaning piglets, and exhibits a wide tissue tropism (Aryal and Liu, 2021). Porcine type-C oncovirus species, which accounted for the highest number of reads in surface samples, consists of Porcine Endogenous Retroviruses (PERVs) and constitute an integral part of the porcine genome (Lopata et al., 2018). Other swine viruses of veterinary interest found in both aerosol and surface samples include porcine circoviruses, parvoviruses and bocaviruses, all of which are associated with porcine respiratory disease complex (PRDC) (Qin et al., 2018). Porcine astroviruses and porcine kobuviruses have been associated with neonatal piglet diarrhea (Qiu et al., 2022). Porcine lymphotropic herpesviruses (PLHVs) are widespread in pigs and, although no association between PLHVs and any pig diseases has been described (Denner (2021); Halecker et al. (2022) suggested that they could be involved in the pathogenesis of erythema multiforme diagnosed in sows. *Sus scrofa* papillomavirus 1 (SsPV1), a member of Dyodeltapapillomavirus 1 species, has previously been described in domestic pigs and has been attributed to papillomatosis tumor (Li et al., 2023; Link et al., 2017).

Additional families, including *Anelloviridae*, *Genomoviridae* and *Tobamoviridae*, were exclusively detected in aerosol samples. Among these viral families, Torque teno sus virus (TTSuV) has also been associated with PRDC (Qin et al., 2018). Porcine torovirus is another potential enteric swine pathogen, and its interspecies recombinant nature with Bovine torovirus has been revealed (Ito et al., 2016). Both toroviruses were detected in aerosol samples. Members of *Polyomaviridae* and *Sedoreoviridae* families, Porcine polyomavirus (PPyV) and Porcine rotavirus A (RVA) respectively, were found only in surface samples. A previous study described PPyV in nasal swabs of pigs with respiratory disease (Hause et al., 2018), while RVA has primarily been associated with diarrhea but also has been identified in pulmonary infections in pigs with respiratory diseases (Nelsen et al., 2022).

All these NGS results provided essential information about the vertebrate viruses to which WWTP and swine farm workers may potentially be exposed. TES can be employed as a tool for WWTPs and farms to screen for the most significant pathogenic viruses present in aerosols and/or surfaces for QMRA. Additionally, it helps identify the most abundant genotypes and assess the emergence of zoonotic viruses on farms or the evolution of animal viruses, potentially increasing the zoonotic risk.

The QMRA analyses conducted in this study aimed to estimate the occupational risk of WWTP workers' exposure to aerosols and contaminated surfaces during their work-related tasks in different seasons of the year. These analyses revealed a noteworthy risk of illness for WWTP workers if safety measures are not taken, due to the inhalation and oral ingestion of HAdV through exposure to workplace bioaerosols and contaminated surfaces. In contrast, the QMRA analysis in the swine farm is intended to be a simulation, as a hypothetical scenario that considers the presence of a virus with zoonotic potential and characteristics similar to PAdV. This extrapolation suggested a possible risk to swine farm workers if all these factors occur and if safety measures are not implemented. These findings align with previous epidemiological studies that have reported work-related symptoms and health effects among WWTP (Al-Batanony and El-Shafie, 2011; Douwes et al., 2001; Thorn and Kerekes, 2001) and swine farms workers (Andersen et al., 2004; Donham et al., 1995; Samadi et al., 2013). These studies further support the association between occupational hazards in these workplaces and adverse health outcomes.

In the WWTP setting, the daily probability of illness due to HAdV inhalation was found to be higher during summer compared to winter, while the higher risk resulting from HAdV oral ingestion was observed in the workspaces during winter. This difference in risk can be attributed to the variations in the concentrations of HAdV in the WWTP aerosols and surfaces explained before. The mean seasonal probability of illness for a WWTP worker exceeded 90% for both transmission routes in both seasons, except for oral ingestion from contaminated surfaces in the break room and inhalation during winter. All these seasonal probabilities would be exceeding the U.S. EPA benchmark ($\leq 10E-4$ pppy). Inhalation of aerosols produced in a WWTP has been suggested as the primary exposure pathway for WWTP workers (Hsiao et al., 2020) and several studies have conducted QMRA to determine the risk of illness caused by inhalation exposure to HAdV (Carducci et al., 2016, 2018), and SARS-CoV-2 (Dada and Gyawali, 2021; Gholipour et al., 2021). The QMRA analyses by Carducci and colleagues also indicated a high-risk of illness for wastewater workers due to exposure to bioaerosols, reporting a higher average risk in sewage influent and biological oxidation tanks (15.64% and 12.73% for an exposure of 3 min) (Carducci et al., 2018). On the other hand, the study conducted by Amoah and coworkers reported that hand-to-mouth ingestion was the major route of exposure during untreated wastewater exposition at the head of the works (Amoah et al., 2022).

In the swine farm, a hypothetical virus with zoonotic potential and similar characteristics to PAdV was chosen for the model. It is important to note that PAdV is a virus that infects pigs and not humans, and the QMRA analysis conducted in the swine farm scenario approximates the

risk of workers getting infected with a hypothetical zoonotic virus. The highest probability of illness occurred through this hypothetical virus oral ingestion from contaminated surfaces, especially during the summer season. The mean seasonal probability of illness for a swine farm worker was 100% for both transmission routes in both seasons, exceeding the U.S. EPA benchmark of $\leq 10E-4$ pppy. A study assessed the occupational risk of zoonotic influenza infection in swine workers, and it was found that spending 25 min working in a barn during an influenza outbreak in a swine herd could be sufficient to cause zoonotic infection in a worker through airborne transmission (Pachha et al., 2016).

The importance of the input variables in QMRA analysis was identified through a sensitivity analysis, which tests the relative impact of stochastic input variables on the result of the models (Federigi et al., 2019). In the aerosol model, the analysis revealed that the parameter with the greatest influence on the model's output, P_{in} variations, was the inhalation rate (f_{in}) in WWTP scenarios and the concentration of PAdV (C_{PAdV}) in swine farm. HAdV concentration was the predominant factor in the sensitivity analysis in the estimated risk described by Carducci et al. (2018). In the surface model, the parameter with the highest influence was the number of hand-to-mouth contacts ($f_{cont(mouth)}$) in nearly all scenarios for both WWTP and the swine farm, except for the break room in the WWTP during both seasons, where the concentration of HAdV (C_{HAdV}) parameter had the greatest influence. These results are consistent with the sensitivity analysis performed by Lanzarini et al. (2022), which reported that the most impacting parameters in the hand-to-mouth model were the frequency of hand-to-mouth contact and the concentration of HAdV.

While QMRA is a valuable tool for estimating occupational risk, it is essential to acknowledge the limitations of the analysis. Firstly, although this study is focused on HAdV and PAdV, the QMRA analysis could potentially be applied to other biological agents that pose an occupational risk for WWTP and swine farm workers. However, selecting adenoviruses as an index pathogen due to their prevalence and environmental persistence allows for a conservative estimation of the levels of other pathogens in the environment (Carducci et al., 2018). It is important to note that PAdV is a virus that infects pigs and not humans. It was chosen as a model to approximate the risk of workers getting infected with a hypothetical virus assumed to have zoonotic potential and infectivity equivalent to adenovirus types. Nevertheless, the probability of spillover for this virus was not incorporated into the models. Additionally, the concentration of the viral pathogens was determined using qPCR, but the infectivity potential was not verified. Therefore, the assumption of a fixed ratio of 700 GC = 1 TCID₅₀ across all scenarios in this study might not accurately reflect the true infectivity potential. Furthermore, the dose-response models used from Teunis et al. (2016) were applied indiscriminately to HAdV and PAdV, even though the study focused on AdV-4, AdV-7, and AdV-16 adenovirus types. Important factors such as secondary transmission of infection, immunity, and variations among workers based on gender, age or health conditions were not considered in the analysis. The deposition efficiency of aerosols in the respiratory tract was also not contemplated in the model for the inhalation route. These limitations highlight the need for further investigation in future studies to enhance the accuracy of QMRA.

The QMRA and NGS analysis performed in this study will facilitate evidence-based decision-making by managers and provide new insights into protection measures and good practices for workers. Assessing occupational risk is essential, and the measures to control it must be defined, as established by the European Union (EU) Directive 2000/54/EC on the protection of workers in the case of activities involving exposure to biological agents (European Commission, 2000). Nevertheless, it is important to note that the directive does not specify exposure limits. The application of QMRA can provide valuable insights for the decision-making process in occupational risk management. By understanding the risks associated with microbial hazards in the workplace, decision-makers can effectively identify and prioritize these risks.

This enables the development of appropriate management strategies and the implementation of safety measures to protect workers.

From a risk management perspective, controlling key factors such as pathogen concentration and exposure time can help mitigate the risks associated with exposure to biological agents. Additionally, increasing the frequency of professional trainings on workplace hazards, implementing protective measures for workers like masks and gloves, and enforcing mandatory vaccinations among WWTP workers can reduce potential risk effects (Jaremków and Agata Kawalec, 2018). The use of personal protective equipment (PPE), including respirators and eye protection, is recommended for swine workers when working with sick animal to reduce risk of cross-species transmission of IAV as well as exposure to other contaminants on swine facilities (Paccha et al., 2016).

These measures collectively contribute to a safer working environment and better protection for workers. The aerosol and surface sampling methods used in the present study could be employed in swine farm as a noninvasive and efficient means to detect and characterize circulating virus in these facilities, as previously suggested for conducting surveillance of novel influenza viruses and other animal viruses through bioaerosol sampling (Anderson et al., 2016). Other pathogenic swine viruses of veterinary interest that could be monitored are PRRSV, PEDV, CSF and ASFV.

Detecting specific viruses in both swine farm aerosol and surface samples, specially through NGS approaches, could serve as an early warning method for the detection of emerging zoonotic pathogens.

5. Conclusions

WWTP and swine farm workers are exposed to viruses during their workday and could face occupational risks associated with exposure to viral pathogens if measures are not implemented. Viral contamination of aerosol and surface samples from these workplaces was identified by qPCR assays, and vertebrate pathogenic viruses were detected through TES.

The application of QMRA models to evaluate the risks associated with aerosol and surface exposures provides valuable estimations of the occupational risk in these workplaces. The daily illness probability due to HAdV inhalation in WWTP was higher in summer, whereas the greatest risk for HAdV oral ingestion from contaminated surfaces was observed in the workspaces during winter. In the swine farm QMRA simulation, considering a hypothetical virus with zoonotic potential and characteristic similar to PAdV, the highest probability of illness occurred through oral ingestion from contaminated surfaces, especially during the summer season.

Overall, this study highlights the significance of evaluating and managing viral pathogen exposure risks in occupational settings to protect the health and well-being of WWTP and swine farm workers. To minimize the risk of virus transmission, it is advisable to implement effective cleaning procedures capable of degrading viral particles on frequently touched surfaces and objects. Additionally, the use of personal protective equipment, especially in areas where bioaerosol particles are aerosolized and can pose a risk, should be made compulsory. Aerosol and surface sampling of WWTP and swine facilities could be implemented as monitoring tools for conducting surveillance to detect emerging zoonotic pathogens.

Data availability statement

The raw sequencing data generated during the current study are available in Zenodo under the DOI number 10.5281/zenodo.8424685.

Patents

The use of paper-based stickers used in this study as a sampling method for long-term monitoring of surfaces is disclosed on Patent Application Publication No. WO 2020/182,924 A1, titled 'A Novel

Sampling Method for Long-Term Monitoring of Microbes' (<https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2020182924>).

The inventors of this patent application are Martin Bobal, Anna Witte, Patrick Mester, and Peter Rossmanith, and the current applicant is Merck Patent GmbH (Frankfurter Strasse 250, 64,293 Darmstadt).

CRediT authorship contribution statement

Marta Itarte: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation. **Miquel Calvo:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Software, Methodology, Formal analysis, Data curation. **Lola Martínez-Frago:** Methodology, Investigation. **Cristina Mejías-Molina:** Writing – review & editing, Visualization, Investigation. **Sandra Martínez-Puchol:** Writing – review & editing, Visualization, Methodology, Investigation. **Rosina Girones:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Gertjan Medema:** Writing – review & editing, Validation, Supervision, Formal analysis. **Sílvia Bofill-Mas:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Marta Rusiñol:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Investigation, Data curation, Conceptualization.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

The authors would like to extend their gratitude to the staff of the WWTP and swine farm involved in this study for their assistance and technical support during the sampling campaigns. Additionally, special appreciation goes to Consell Comarcal del Maresme and SIMMAR for their kind assistance in the WWTP sampling, as well as Vicenç, Jordi Baliellas and GSP in the swine farm.

This research was conducted as part of the EPISARS project (544/C/2021), which received funding from donations obtained during “La Marató de TV3” dedicated to COVID-19, and the VIRALERT project (PID 2021-128014OB-I00), funded by MCIN/AEI/10.13039/501100011033/FEDER, UE. Sílvia Bofill-Mas is a Serra-Hunter fellow at the University of Barcelona.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijheh.2024.114360>.

References

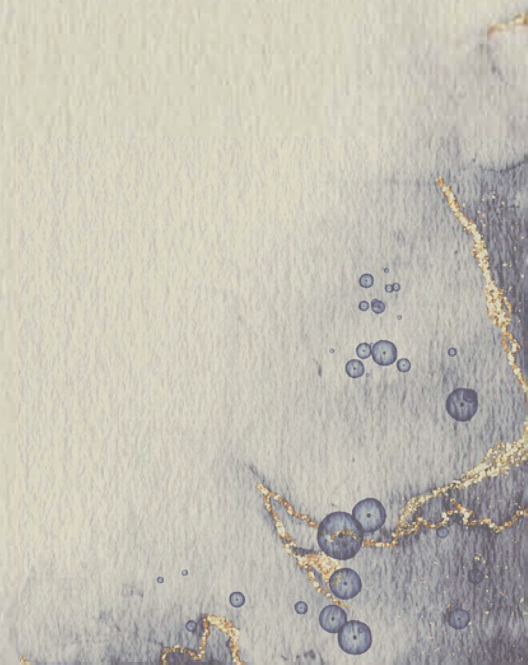
- Agranovski, V., Ristovski, Z., Blackall, P.J., Morawska, L., 2004. Size-selective assessment of airborne particles in swine confinement building with the UVAPS. *Atmos. Environ.* 38, 3893–3901. <https://doi.org/10.1016/j.atmosenv.2004.02.058>.
- Al-Batanony, M.A., El-Shafie, M.K., 2011. Work-related health effects among wastewater treatment plants workers. *Int. J. Occup. Environ. Med.* 2, 237–244.
- Allard, A., Vantarakis, A., 2017. Adenoviruses. *Glob. Water pathog. Proj.* <https://doi.org/10.14321/waterpathogens.11>.
- Alonso, G., Raynor, P.C., Davies, P.R., Torremorell, M., 2015. Concentration, size distribution, and infectivity of airborne particles carrying swine viruses. *PLoS One* 10, 1–12. <https://doi.org/10.1371/journal.pone.0135675>.
- Amoah, I.D., Kumari, S., Bux, F., 2022. A probabilistic assessment of microbial infection risks due to occupational exposure to wastewater in a conventional activated sludge wastewater treatment plant. *Sci. Total Environ.* 843, 156849 <https://doi.org/10.1016/j.scitotenv.2022.156849>.
- Andersen, C.I., Von Essen, S.G., Smith, L.M., Spencer, J., Jolie, R., Donham, K.J., 2004. Respiratory symptoms and airway obstruction in swine veterinarians: a persistent problem. *Am. J. Ind. Med.* 46, 386–392. <https://doi.org/10.1002/ajim.20080>.
- Anderson, B.D., Ma, M., Xia, Y., Wang, T., Shu, B., Lednický, J.A., Ma, M.J., Lu, J., Gray, G.C., 2016. Bioaerosol sampling in modern agriculture: a novel approach for

- emerging pathogen surveillance? *J. Infect. Dis.* 214, 537–545. <https://doi.org/10.1093/infdis/jiw180>.
- Anderson, B.D., Yondon, M., Bailey, E.S., Duman, E.K., Simmons, R.A., Greer, A.G., Gray, G.C., 2021. Environmental bioaerosol surveillance as an early warning system for pathogen detection in North Carolina swine farms: a pilot study. *Transbound. Emerg. Dis.* 68, 361–367. <https://doi.org/10.1111/tbed.13683>.
- Antwi-Agyei, P., Biran, A., Peasey, A., Bruce, J., Ensink, J., 2016. A faecal exposure assessment of farm workers in Accra, Ghana: a cross sectional study. *BMC Publ. Health* 16, 1–13. <https://doi.org/10.1186/s12889-016-3266-8>.
- Aryal, M., Liu, G., 2021. Porcine bocavirus: a 10-year history since its discovery. *Virology* 36, 1261–1272. <https://doi.org/10.1007/s12250-021-00365-z>.
- Benfield, D.A.H., Richard, R.A., 2012. Porcine adenoviruses. In: Zimmerman, J.J., Karriker, L.A., Ramirez, A., Schwartz, K.J., Stevenson, G.W. (Eds.), *Diseases of Swine*, tenth ed. Wiley, New Jersey, pp. 392–395.
- Bobal, M., Witte, A.K., Mester, P., Fister, S., Schoder, D., Rossmann, P., 2019. A novel method for sampling and long-term monitoring of Microbes that uses stickers of plain paper. *Appl. Environ. Microbiol.* 85, 1–11. <https://doi.org/10.1128/AEM.00766-19>.
- Bofill-Mas, S., Albinana-Gimenez, N., Clemente-Casares, P., Hunders, A., Rodriguez-Manzano, J., Allard, A., Calvo, M., Girones, R., 2006. Quantification and stability of human adenoviruses and polyomavirus JCPyV in wastewater matrices. *Appl. Environ. Microbiol.* 72, 7894–7896. <https://doi.org/10.1128/AEM.00965-06>.
- Bofill-Mas, S., Rusiñol, M., Fernandez-Cassi, X., Carratalà, A., Hunders, A., Girones, R., 2013. Quantification of human and animal viruses to differentiate the origin of the fecal contamination present in environmental samples. *BioMed Res. Int.* <https://doi.org/10.1155/2013/192089>, 2013.
- Boucherbaine, S., Nassar, R., Zaher, S., Mohamed, L., Olsen, M., Alqutami, F., Hachim, M., Alkhaja, A., Campos, M., Jones, P., McKirdy, S., Alghafri, R., Tajouri, L., Senok, A., 2022. Metagenomic sequencing and reverse transcriptase PCR reveal that mobile phones and environmental surfaces are reservoirs of multidrug-resistant superbugs and SARS-CoV-2. *Front. Cell. Infect. Microbiol.* 12, 1–11. <https://doi.org/10.3389/fcimb.2022.806077>.
- Briesse, T., Kapoor, A., Mishra, N., Jain, K., Kumar, A., Jabado, O.J., Lipkin, W.I., 2015. Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *mBio* 6. <https://doi.org/10.1128/mBio.01491-15>.
- Brisebois, E., Veillette, M., Dion-Dupont, V., Lavoie, J., Corbeil, J., Culley, A., Duchaine, C., 2018. Human viral pathogens are pervasive in wastewater treatment center aerosols. *J. Environ. Sci.* 67, 45–53. <https://doi.org/10.1016/j.jes.2017.07.015>.
- Brooks, J.P., Tanner, B.D., Josephson, K.L., Gerba, C.P., Haas, C.N., Pepper, I.L., 2005. A national study on the residential impact of biological aerosols from the land application of biosolids. *J. Appl. Microbiol.* 99, 310–322. <https://doi.org/10.1111/j.1365-2672.2005.02604.x>.
- Brooks, J.P., McLaughlin, M.R., Gerba, C.P., Pepper, I.L., 2012. Land application of manure and class B biosolids: an occupational and public quantitative microbial risk assessment. *J. Environ. Qual.* 41, 2009–2023. <https://doi.org/10.2134/jeq2011.0430>.
- Canales, R.A., Wilson, A.M., Pearce-Walker, J.I., Verhoughstraete, M.P., Reynolds, K.A., 2018. Methods for handling left-censored data in quantitative microbial risk assessment. *Appl. Environ. Microbiol.* 84. <https://doi.org/10.1128/AEM.01203-18>.
- Carducci, A., Donzelli, G., Cioni, L., Verani, M., 2016. Quantitative microbial risk assessment in occupational settings applied to the airborne human adenovirus infection. *Int. J. Environ. Res. Publ. Health* 13. <https://doi.org/10.3390/ijerph13070733>.
- Carducci, A., Donzelli, G., Cioni, L., Federigi, I., Lombardi, R., Verani, M., 2018. Quantitative microbial risk assessment for workers exposed to bioaerosol in wastewater treatment plants aimed at the choice and setup of safety measures. *Int. J. Environ. Res. Publ. Health* 15. <https://doi.org/10.3390/ijerph15071490>.
- Carratalà, A., Rusiñol, M., Hunders, A., Biarnes, M., Rodriguez-Manzano, J., Vantarakis, A., Kern, A., Suñen, E., Girones, R., Bofill-Mas, S., 2012. A novel tool for specific detection and quantification of chicken/Turkey parvoviruses to trace poultry fecal contamination in the environment. *Appl. Environ. Microbiol.* 78, 7496–7499. <https://doi.org/10.1128/AEM.01283-12>.
- CDC Division of Viral Diseases, 2023. 2019–Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel: Instructions for Use. CDC-006-00019, Revis. 08.
- Chen, Y., Yan, C., Yang, Y., Ma, J., 2021. Quantitative microbial risk assessment and sensitivity analysis for workers exposed to pathogenic bacterial bioaerosols under various aerated modes in two wastewater treatment plants. *Sci. Total Environ.* 755.
- Cherrie, J.W., Semple, S., Christopher, Y., Saleem, A., Hughson, G.W., Phillips, A., 2006. How important is inadvertent ingestion of hazardous substances at work? *Ann. Occup. Hyg.* 50, 693–704. <https://doi.org/10.1093/annhyg/mel035>.
- Corpus, M.V.A., Buonerba, A., Vigliotta, G., Zarra, T., Ballesteros, F., Campiglia, P., Belgioino, V., Korshin, G., Nadeo, V., 2020. Viruses in wastewater: occurrence, abundance and detection methods. *Sci. Total Environ.* 745. <https://doi.org/10.1016/j.scitotenv.2020.140910>.
- Corrao, C.R.N., Mazzotta, A., La Torre, G., De Giusti, M., 2012. Biological risk and occupational Health. *Ind. Health* 50, 326–337. <https://doi.org/10.2486/indhealth.MS1324>.
- Corzo, C.A., Culhane, M., Dee, S., Morrison, R.B., Torremorell, M., 2013. Airborne detection and quantification of swine influenza A virus in air samples collected inside, outside and downwind from swine barns. *PLoS One* 8, 1–7. <https://doi.org/10.1371/journal.pone.0071444>.
- Dada, A.C., Gyawali, P., 2021. Quantitative microbial risk assessment (QMRA) of occupational exposure to SARS-CoV-2 in wastewater treatment plants. *Sci. Total Environ.* 763. <https://doi.org/10.1016/j.scitotenv.2021.142989>.
- De Motes, C.M., Clemente-Casares, P., Hunders, A., Martín, M., Girones, R., 2004. Detection of bovine and porcine adenoviruses for tracing the source of fecal contamination. *Appl. Environ. Microbiol.* 70, 1448–1454. <https://doi.org/10.1128/AEM.70.3.1448-1454.2004>.
- Dee, S., Otake, S., Oliveira, S., Deen, J., 2009. Evidence of long distance airborne transport of porcine reproductive and respiratory syndrome virus and Mycoplasma hyopneumoniae. *Vet. Res.* 40. <https://doi.org/10.1051/vetres/2009022>.
- Deere, D., Ryan, U., 2022. Current assumptions for quantitative microbial risk assessment (QMRA) of Norovirus contamination of drinking water catchments due to recreational activities: an update. *J. Water Health* 20, 1543–1557. <https://doi.org/10.2166/wh.2022.114>.
- Denner, J., 2021. Porcine lymphotropic herpesviruses (PLHVS) and xenotransplantation. *Viruses* 13. <https://doi.org/10.3390/v13061072>.
- Di Bonito, P., Della Libera, S., Petrica, S., Iaconelli, M., Sanguinetti, M., Graffeo, R., Accardi, L., La Rosa, G., 2015. A large spectrum of alpha and beta papillomaviruses are detected in human stool samples. *J. Gen. Virol.* 96, 607–613. <https://doi.org/10.1099/vir.0.071787-0>.
- Donham, K.J., Reynolds, S.J., Whitten, P., Merchant, J.A., Burmeister, L.F., Popenorf, W.J., 1995. Respiratory dysfunction in swine production facility workers: dose-response relationships of environmental exposures and pulmonary function. *Am. J. Ind. Med.* 27, 405–418.
- Douwes, J., Mannetje, A., Heederik, D., 2001. Work-related symptoms in sewage treatment workers. *Ann. Agric. Environ. Med.* 8, 39–45.
- Douwes, J., Thorne, P., Pearce, N., Heederik, D., 2003. Bioaerosol health effects and exposure assessment: progress and prospects. *Ann. Occup. Hyg.* 47, 187–200. <https://doi.org/10.1093/annhyg/mec032>.
- Duhaime, M.B., Deng, L., Poulos, B.T., Sullivan, M.B., 2012. Towards quantitative metagenomics of wild viruses and other ultra-low concentration DNA samples: a rigorous assessment and optimization of the linker amplification method. *Environ. Microbiol.* 14, 2526–2537. <https://doi.org/10.1111/j.1462-2920.2012.02791.x>.
- Egawa, N., Egawa, K., Griffin, H., Doorbar, J., 2015. Human papillomaviruses; Epithelial tropisms, and the development of neoplasia. *Viruses* 7, 3863–3890. <https://doi.org/10.3390/v7072802>.
- EPA, 2011. Exposure factors handbook: chapter 6—inhalation rates [WWW Document]. URL: <https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=236252>. (Accessed 10 February 2023).
- European Commission, 2000. Directive 2000/54/EC of the European parliament and of the council of 18 september 2000 on the protection of workers from risks related to exposure to biological agents at work (seventh individual directive within the meaning of article 16(1) of directive [WWW Document]. URL: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32000L0054&from=EN>. (Accessed 10 February 2023).
- Federigi, I., Verani, M., Donzelli, G., Cioni, L., Carducci, A., 2019. The application of quantitative microbial risk assessment to natural recreational waters: a review. *Mar. Pollut. Bull.* 144, 334–350. <https://doi.org/10.1016/j.marpolbul.2019.04.073>.
- Fernandez-Cassi, X., Rusiñol, M., Martínez-Puchol, S., 2018. Viral concentration and amplification from human serum samples prior to application of next-generation sequencing analysis. *Methods Mol. Biol.* 173–188.
- Fernandez-Cassi, X., Timoneda, N., Martínez-Puchol, S., Rusiñol, M., Rodriguez-Manzano, J., Figuerola, N., Bofill-Mas, S., Abril, J.F., Girones, R., 2018. Metagenomics for the study of viruses in urban sewage as a tool for public health surveillance. *Sci. Total Environ.* 618, 870–880. <https://doi.org/10.1016/j.scitotenv.2017.08.249>.
- Filipa-Silva, A., Parreira, R., Martínez-Puchol, S., Bofill-Mas, S., Barreto Crespo, M.T., Nunes, M., 2020. The unexplored virome of two atlantic coast fish: contribution of next-generation sequencing to fish virology. *Foods* 9, 1634. <https://doi.org/10.3390/foods9111634>.
- Forés, E., Bofill-Mas, S., Itarte, M., Martínez-Puchol, S., Hunders, A., Calvo, M., Borrego, C.M., Corominas, L.L., Girones, R., Rusiñol, M., 2021. Evaluation of two rapid ultrafiltration-based methods for SARS-CoV-2 concentration from wastewater. *Sci. Total Environ.* 768. <https://doi.org/10.1016/j.scitotenv.2020.144786>.
- Frazzini, S., Amadori, M., Turin, L., Riva, F., 2022. SARS CoV-2 infections in animals, two years into the pandemic. *Arch. Virol.* 167, 2503–2517. <https://doi.org/10.1007/s00705-022-05609-1>.
- Gholipour, S., Mohammadi, F., Nikaeen, M., Shamsizadeh, Z., Khazeni, A., Sahbaei, Z., Mousavi, S.M., Ghobadian, M., Mirhendi, H., 2021. COVID-19 infection risk from exposure to aerosols of wastewater treatment plants. *Chemosphere* 273, 129701. <https://doi.org/10.1016/j.chemosphere.2021.129701>.
- Gray, G.C., McCarthy, T., Capuano, A.W., Setterquist, S.F., Olsen, C.W., Alavanja, M.C., Lynch, C.F., 2007. Swine workers and swine influenza virus infections. *Emerg. Infect. Dis.* 13, 1871–1878. <https://doi.org/10.3201/eid1312.061323>.
- Grisoli, P., Rodolfi, M., Villani, S., Grignani, E., Cottica, D., Berri, A., Maria Picco, A., Dacarro, C., 2009. Assessment of airborne microorganism contamination in an industrial area characterized by an open composting facility and a wastewater treatment plant. *Environ. Res.* 109, 135–142. <https://doi.org/10.1016/j.envres.2008.11.001>.
- Haas, C.N., Rose, J.B., Gerba, C.P., 1999. *Quantitative Microbial Risk Assessment*. John Wiley and Sons Inc, New York, NY, USA.
- Halecker, S., Papatsiros, V., Psalla, D., Krabben, L., Kaufer, B., Denner, J., 2022. Virological characterization of pigs with erythema multiforme. *Microorganisms* 10, 1–10. <https://doi.org/10.3390/microorganisms10030652>.
- Hall, R.J., Leblanc-Maridor, M., Wang, J., Ren, X., Moore, N.E., Brooks, C.R., Peasey, M., Douwes, J., McLean, D.J., 2013. Metagenomic detection of viruses in aerosol samples from workers in animal slaughterhouses. *PLoS One* 8. <https://doi.org/10.1371/journal.pone.0072226>.

- Han, Y., Li, L., Liu, J., 2013. Characterization of the airborne bacteria community at different distances from the rotating brushes in a wastewater treatment plant by 16S rRNA gene clone libraries. *J. Environ. Sci.* 25, 5–15. [https://doi.org/10.1016/S1001-0742\(12\)60018-7](https://doi.org/10.1016/S1001-0742(12)60018-7).
- Han, Y., Yang, T., Chen, T., Li, L., Liu, J., 2019. Characteristics of submicron aerosols produced during aeration in wastewater treatment. *Sci. Total Environ.* 696, 134019. <https://doi.org/10.1016/j.scitotenv.2019.134019>.
- Hause, B.M., Smith, C., Bishop, B., Stewart, C., Simonson, R., 2018. Complete genome sequence of a porcine polyomavirus from nasal swabs of pigs with respiratory disease. *Genome Announc.* 6. <https://doi.org/10.1128/genomeA.00344-18>.
- Hernroth, B.E., Conden-Hansson, A.C., Rehnstam-Holm, A.S., Girones, R., Allard, A.K., 2002. Environmental factors influencing human viral pathogens and their potential indicator organisms in the blue mussel, *Mytilus edulis*: the first Scandinavian report. *Appl. Environ. Microbiol.* 68, 4523–4533. <https://doi.org/10.1128/AEM.68.9.4523-4533.2002>.
- Hijnen, W.A.M., Beerenndonk, E.F., Medema, G.J., 2006. Inactivation credit of UV radiation for viruses, bacteria and protozoan (oo)cysts in water: a review. *Water Res.* 40, 3–22. <https://doi.org/10.1016/j.watres.2005.10.030>.
- Hjelmso, M.H., Mollerup, S., Jensen, R.H., Pietroni, C., Lukjancenko, O., Schultz, A.C., Aarestrup, F.M., Hansen, A.J., 2019. Metagenomic analysis of viruses in toilet waste from long distance flights—a new procedure for global infectious disease surveillance. *PLoS One* 14. <https://doi.org/10.1371/journal.pone.0210368>.
- HSE, 2007. *Inadvertent ingestion exposure in the workplace*, in: Literature Review. In: Christopher, Y., Semple, S., Hughson, G.W., Cherrie, J.W. (Eds.), *Health and Safety Executive. Institute of Occupational Medicine, Research Park North, Riccarton Edinburgh*. EH14 4AP.
- Hsiao, T.C., Lin, A.Y.C., Lien, W.C., Lin, Y.C., 2020. Size distribution, biological characteristics and emerging contaminants of aerosols emitted from an urban wastewater treatment plant. *J. Hazard Mater.* 388, 121809. <https://doi.org/10.1016/j.jhazmat.2019.02.089>.
- Huang, W., Bai, L., Tang, H., 2023. Epstein-Barr virus infection: the micro and macro worlds. *Virol. J.* 20. <https://doi.org/10.1186/s12985-023-02187-9>.
- Hundesda, A., Maluquer De Motes, C., Bofill-Mas, S., Albinana-Gimenez, N., Girones, R., 2006. Identification of human and animal adenoviruses and polyomaviruses for determination of sources of fecal contamination in the environment. *Appl. Environ. Microbiol.* 72, 7886–7893. <https://doi.org/10.1128/AEM.01090-06>.
- Hundesda, A., Maluquer de Motes, C., Albinana-Gimenez, N., Rodriguez-Manzano, J., Bofill-Mas, S., Suñen, E., Rosina Girones, R., 2009. Development of a qPCR assay for the quantification of porcine adenoviruses as an MST tool for swine fecal contamination in the environment. *J. Virol. Methods* 158, 130–135. <https://doi.org/10.1016/j.jviromet.2009.02.006>.
- Iaconelli, M., Petricea, S., Della Libera, S., Di Bonito, P., La Rosa, G., 2015. First detection of human papillomaviruses and human polyomaviruses in river waters in Italy. *Food Environ. Virol.* 7, 309–315. <https://doi.org/10.1007/s12560-015-9203-7>.
- Itarte, M., Martínez-Puchol, S., Forés, E., Hundesa, A., Timoneda, N., Bofill-Mas, S., Girones, R., Rusinöl, M., 2021. NGS techniques reveal a high diversity of RNA viral pathogens and papillomaviruses in fresh produce and irrigation water. *Foods* 10. <https://doi.org/10.3390/foods10081820>.
- Ito, M., Tsuchiaka, S., Naoi, Y., Otomaru, K., Sato, M., Masuda, T., et al., 2016. Whole genome analysis of Japanese bovine toroviruses reveals natural recombination between porcine and bovine toroviruses. *Infect Genet Evol* 38, 90–95. <https://doi.org/10.1016/j.meegid.2015.12.013>.
- Jaremków, Aleksandra, Agata Kawalec, K.P., 2018. Harmful factors in wastewater treatment plant - knowledge and awareness of workers about hazards. *Probl. Hig. Epidemiol.* 99, 189–195.
- Kalantar, K.L., Carvalho, T., De Bourcy, C.F.A., Dimitrov, B., Dingle, G., Egger, R., Han, J., Holmes, O.B., Juan, Y.F., King, R., Kislyuk, A., Lin, M.F., Mariano, M., Morse, T., Reynoso, L.V., Cruz, D.R., Sheu, J., Tang, J., Wang, J., Zhang, M.A., Zhong, E., Ahyong, V., Lay, S., Chea, S., Bohl, J.A., Manning, J.E., Tato, C.M., DeRisi, J.L., 2020. IDseq-An open source cloud-based pipeline and analysis service for metagenomic pathogen detection and monitoring. *GigaScience* 9, 1–14. <https://doi.org/10.1093/gigascience/giaa111>.
- Karavarsamis, N., Hamilton, A.J., 2010. Estimators of annual probability of infection for quantitative microbial risk assessment. *J. Water Health* 8, 365–373. <https://doi.org/10.2166/wh.2010.045>.
- Karlsson, O.E., Belák, S., Granberg, F., 2013. The Effect of Preprocessing by Sequence-independent, Single-Primer Amplification (SISPA) on Metagenomic Detection of Viruses, vol. 11, pp. S227–S234. <https://doi.org/10.1089/bsp.2013.0008>.
- Kataki, S., Patowary, R., Chatterjee, S., Vairale, M.G., Sharma, S., Dwivedi, S.K., Kamboj, D.V., 2022. Bioaerosolization and pathogen transmission in wastewater treatment plants: microbial composition, emission rate, factors affecting and control measures. *Chemosphere* 287, 132180. <https://doi.org/10.1016/j.chemosphere.2021.132180>.
- Kumthip, K., Khamrin, P., Kongkaew, A., Vachirachewin, R., Malasao, R., Ushijima, H., Maneekarn, N., 2019. Molecular epidemiology and characterization of porcine adenoviruses in pigs with diarrhea in Thailand. *Infect. Genet. Evol.* 67, 73–77. <https://doi.org/10.1016/j.meegid.2018.10.026>.
- La Rosa, G., Frattini, M., Accardi, L., D'Oro, G., Della Libera, S., Muscillo, M., Di Bonito, P., 2013. Mucosal and cutaneous human papillomaviruses detected in raw sewage. *PLoS One* 8, e52391. <https://doi.org/10.1371/journal.pone.0052391>.
- Lanzarini, N.M., Federigi, I., Marinho Mata, R., Neves Borges, M.D., Mendes Saggiore, E., Cioni, L., Verani, M., Carducci, A., Costa Moreira, J., Ferreira Mannarino, C., Pereira Miagostovich, M., 2022. Human adenovirus in municipal solid waste leachate and quantitative risk assessment of gastrointestinal illness to waste collectors. *Waste Manag.* 138, 308–317. <https://doi.org/10.1016/j.wasman.2021.12.005>.
- Li, Y., Zhang, X., Zhao, C., Lei, X., Huang, H., Shi, Y., Li, C., Bi, J., Sun, W., Lan, T., Zheng, M., 2023. Genetic characterization of Sus scrofa papillomavirus type 1 from domestic pigs in Guangxi Province, China. *Braz. J. Microbiol.* 54, 2437–2443. <https://doi.org/10.1007/s42770-023-01092-1>.
- Lim, K.Y., Hamilton, A.J., Jiang, S.C., 2015. Assessment of public health risk associated with viral contamination in harvested urban stormwater for domestic applications. *Sci. Total Environ.* 523, 95–108. <https://doi.org/10.1016/j.scitotenv.2015.03.077>.
- Link, E.K., Hoferer, M., Strohe, B., Rigbers, K., Langenmayer, M.C., Sutter, G., Fux, R., 2017. Sus scrofa papillomavirus 2-genetic characterization of a novel suid papillomavirus from wild boar in Germany. *J. Gen. Virol.* 98, 2113–2117. <https://doi.org/10.1099/jgv.0.000868>.
- Liu, W., MacDonald, M., You, J., 2016. Merkel cell polyomavirus infection and Merkel cell carcinoma. *Curr. Opin. Virol.* 20, 20–27. <https://doi.org/10.1016/j.coviro.2016.07.011>.
- Liu, W., Krump, N.A., MacDonald, M., You, J., 2018. Merkel cell polyomavirus infection of animal dermal fibroblasts. *J. Virol.* 92, 1–14. <https://doi.org/10.1128/jvi.01610-17>.
- Liu, T., Li, G., Liu, Z., Xi, L., Ma, W., Gao, X., 2023. Characteristics of aerosols from swine farms: a review of the past two-decade progress. *Environ. Int.* 178, 108074. <https://doi.org/10.1016/j.envint.2023.108074>.
- Lopata, K., Wojdas, E., Nowak, R., Lopata, P., Mazurek, U., 2018. Porcine endogenous retrovirus (PERV)-molecular structure and replication strategy in the context of retroviral infection risk of human cells. *Front. Microbiol.* 9, 1–11. <https://doi.org/10.3389/fmicb.2018.00730>.
- Lopez, G.U., Gerba, C.P., Tamimi, A.H., Kitajima, M., Maxwell, S.L., Rose, J.B., 2013. Transfer efficiency of bacteria and viruses from porous and nonporous fomites to fingers under different relative humidity conditions. *Appl. Environ. Microbiol.* 79, 5728–5734. <https://doi.org/10.1128/AEM.01030-13>.
- López-Lorenzo, G., López-Navo, C., Prieto, A., Díaz, P., Panadero, R., Rodríguez-Vega, V., Morondo, P., Fernández, G., Díaz-Cao, J.M., 2022. Monitoring of porcine coronavirus type 2 infection through air and surface samples in vaccinated and unvaccinated fattening farms. *Transbound. Emerg. Dis.* 69, 1108–1117. <https://doi.org/10.1111/tbed.14069>.
- Mandal, J., Brandl, H., 2011. Bioaerosols in indoor environment - a review with special reference to residential and occupational locations. *Open Environ. Biol. Monit. J.* 4, 83–96. <https://doi.org/10.2174/1875040001104010083>.
- Martínez-Puchol, S., Rusinöl, M., Fernández-Cassi, X., Timoneda, N., Itarte, M., Andrés, C., Antón, A., Abril, J.F., Girones, R., Bofill-Mas, S., 2020. Characterisation of the sewage virome: comparison of NGS tools and occurrence of significant pathogens. *Sci. Total Environ.* 713. <https://doi.org/10.1016/j.scitotenv.2020.136604>.
- Martínez-Puchol, S., Cardona, L., Drago, M., Gazo, M., Bofill-Mas, S., 2022. Viral metagenomics reveals persistent as well as dietary acquired viruses in Antarctic Fur seals. *Sci. Rep.* 12, 1–11. <https://doi.org/10.1038/s41598-022-23114-y>.
- Masclaux, F.G., Hotz, P., Gashi, D., Savova-Bianchi, D., Opplinger, A., 2014. Assessment of airborne virus contamination in wastewater treatment plants. *Environ. Res.* 133, 260–265. <https://doi.org/10.1016/j.envres.2014.06.002>.
- McBride, G.B., Stott, R., Miller, W., Bambic, D., Wuertz, S., 2013. Discharge-based QMRA for estimation of public health risks from exposure to stormwater-borne pathogens in recreational waters in the United States. *Water Res.* 47, 5282–5297. <https://doi.org/10.1016/j.watres.2013.06.001>.
- Medema, G., Wuillings, B., Roelvelde, P., Van Der Kooij, D., 2004. Risk assessment of Legionella and enteric pathogens in sewage treatment works. *Water Sci. Technol. Water Supply* 4, 125–132. <https://doi.org/10.2166/ws.2004.0037>.
- Mejías-Molina, C., Pico-Tomás, A., Beltrán-Rubín, A., Martínez-Puchol, S., Corominas, L., Rusinöl, M., Bofill-Mas, S., 2023. Effectiveness of passive sampling for the detection and genetic characterization of human viruses in wastewater. *Environ. Sci. Water Res. Technol.* <https://doi.org/10.1039/d2ew00867j>.
- Mirskaya, E., Agranovski, I.E., 2018. Sources and mechanisms of bioaerosol generation in occupational environments. *Crit. Rev. Microbiol.* 44, 739–758. <https://doi.org/10.1080/1040841X.2018.1508125>.
- Myers, K.P., Olsen, C.W., Setterquist, S.F., Capuano, A.W., Donham, K.J., Thacker, E.L., Merchant, J.A., Gray, G.C., 2006. Are swine workers in the United States at increased risk of infection with zoonotic influenza virus? *Clin. Infect. Dis.* 42, 14–20. <https://doi.org/10.1086/498977>.
- Neira, V., Rabinowitz, P., Rendahl, A., Paccha, B., Gibbs, S.G., Torremorell, M., 2016. Characterization of viral load, viability and persistence of influenza A virus in air and on surfaces of swine production facilities. *PLoS One* 11, 1–11. <https://doi.org/10.1371/journal.pone.0146616>.
- Nelsen, A., Lager, K.M., Stasko, J., Nelson, E., Lin, C.M., Hause, B.M., 2022. Identification of pulmonary infections with porcine rotavirus A in pigs with respiratory disease. *Front. Vet. Sci.* 9, 1–8. <https://doi.org/10.3389/fvets.2022.918736>.
- Nietfeld, J.C., Leslie-Stein, P., 1993. Interstitial nephritis in pigs with adenovirus infection. *J. Vet. Diagn. Invest.* 5, 269–273. <https://doi.org/10.1177/104063879300500223>.
- Nieuwenhuijsen, D.F., Oude Munnink, B.B., Phan, M.V.T., Munk, P., Venkatakrishnan, S., Aarestrup, F.M., Cotten, M., Koopmans, M.P.G., 2020. Setting a baseline for global urban virome surveillance in sewage. *Sci. Rep.* 10, 1–13. <https://doi.org/10.1038/s41598-020-69869-0>.
- Olesen, A.S., Lohse, L., Boklund, A., Halasa, T., Gallardo, C., Pejsak, Z., Belsham, G.J., Rasmussen, T.B., Botner, A., 2017. Transmission of African swine fever virus from infected pigs by direct contact and aerosol routes. *Vet. Microbiol.* 211, 92–102. <https://doi.org/10.1016/j.vetmic.2017.10.004>.
- Onakpoya, I.J., Heneghan, C.J., Spencer, E.A., Brassey, J., Rosca, E.C., Maltoni, S., Plüddemann, A., Evans, D.H., Conly, J.M., Jefferson, T., 2022. Viral cultures for

- assessing fomite transmission of SARS-CoV-2: a systematic review and meta-analysis. *J. Hosp. Infect.* 130, 63–94. <https://doi.org/10.1016/j.jhin.2022.09.007>.
- One Health Commission, 2018. What is one health? [WWW Document]. URL: https://www.onehealthcommission.org/en/why_one_health/what_is_one_health/. (Accessed 28 September 2023).
- Oppliger, A., Hilfiker, S., Duc, T.V., 2005. Influence of seasons and sampling strategy on assessment of bioaerosols in sewage treatment plants in Switzerland. *Ann. Occup. Hyg.* 49, 393–400. <https://doi.org/10.1093/annhyg/meh108>.
- O'Brien, K.M., Nonnenmann, M.W., 2016. Airborne influenza A is detected in the personal breathing zone of swine veterinarians. *PLoS One* 11, 1–8. <https://doi.org/10.1371/journal.pone.0149083>.
- Paccha, B., Jones, R.M., Gibbs, S., Kane, M.J., Torremorell, M., Neira-Ramirez, V., Rabinowitz, P.M., 2016. Modeling risk of occupational zoonotic influenza infection in swine workers. *J. Occup. Environ. Hyg.* 13, 577–587. <https://doi.org/10.1080/15459624.2016.1159688>.
- Paul, P.S., Halbur, P., Janke, B., Joo, H., Nawagitgul, P., Singh, J., Sorden, S., 2003. Exogenous porcine viruses. *Curr. Top. Microbiol. Immunol.* 278, 125–183. https://doi.org/10.1007/978-3-642-55541-1_6.
- Petterson, S.R., 2016. Application of a QMRA framework to inform selection of drinking water interventions in the developing context. *Risk Anal.* 36, 203–214. <https://doi.org/10.1111/risa.12452>.
- Poh, M.K., Ma, M., Nguyen, T.T., Su, Y.C.F., Pena, E.M., Ogden, B.E., Anderson, B.D., Gray, G.C., 2017. Bioaerosol sampling for airborne respiratory viruses in an experimental medicine pig handling facility, Singapore. *Southeast Asian J. Trop. Med. Publ. Health* 48, 828–835.
- Prost, K., Kloeze, H., Mukhi, S., Bozek, K., Poljak, Z., Mubareka, S., 2019. Bioaerosol and surface sampling for the surveillance of influenza A virus in swine. *Transbound. Emerg. Dis.* 66, 1210–1217. <https://doi.org/10.1111/tbed.13139>.
- Qin, S., Ruan, W., Yue, H., Tang, C., Zhou, K., Zhang, B., 2018. Viral communities associated with porcine respiratory disease complex in intensive commercial farms in Sichuan province, China. *Sci. Rep.* 8, 1–9. <https://doi.org/10.1038/s41598-018-31554-8>.
- Qiu, M., Li, S., Xiao, Y., Li, J., Zhang, Y., Li, X., Feng, B., Li, C., Lin, H., Zhu, J., Chen, N., 2022. Pathogenic and metagenomic evaluations reveal the correlations of porcine epidemic diarrhea virus, porcine kobuvirus and porcine astroviruses with neonatal piglet diarrhea. *Microb. Pathog.* 170, 105703 <https://doi.org/10.1016/j.micpath.2022.105703>.
- R Core, T., 2023. R: A Language and Environment for Statistical Computing [WWW Document]. R Found. Stat. Comput. URL: <https://www.r-project.org/>. (Accessed 28 September 2023).
- Rusin, P., Maxwell, S., Gerba, C., 2002. Comparative surface-to-hand and fingertip-to-mouth transfer efficiency of gram-positive bacteria, gram-negative bacteria, and phage. *J. Appl. Microbiol.* 93, 585–592. <https://doi.org/10.1046/j.1365-2672.2002.01734.x>.
- Rusiñol, M., Girones, R., 2017. Summary of excreted and waterborne viruses. *Glob. Water Pathog. Proj.* <https://doi.org/10.14321/waterpathogens.19>.
- Rusiñol, M., Fernandez-Cassi, X., Hundesa, A., Vieira, C., Kern, A., Eriksson, I., Ziros, P., Kay, D., Miagostovich, M., Vargha, M., Allard, A., Vantarakis, A., Wyn-Jones, P., Boffill-Mas, S., Girones, R., 2014. Application of human and animal viral microbial source tracking tools in fresh and marine waters from five different geographical areas. *Water Res.* 59, 119–129. <https://doi.org/10.1016/j.watres.2014.04.013>.
- Rusiñol, M., Martínez-Puchol, S., Timoneda, N., Fernández-Cassi, X., Pérez-Cataluña, A., Fernández-Bravo, A., Moreno-Mesonero, L., Moreno, Y., Alonso, J.L., Figueras, M.J., Abril, J.F., Boffill-Mas, S., Girones, R., 2020. Metagenomic analysis of viruses, bacteria and protozoa in irrigation water. *Int. J. Hyg. Environ. Health* 224, 113440. <https://doi.org/10.1016/j.ijheh.2019.113440>.
- Salem, E., Häggglund, S., Cassard, H., Corre, T., Näslund, K., Foret, C., Gauthier, D., Pinard, A., Delverdier, M., Zohari, S., Valarcher, J.-F., Ducatez, M., Meyer, G., 2019. Pathogenesis, host innate immune response, and aerosol transmission of influenza D virus in cattle. *J. Virol.* 93 <https://doi.org/10.1128/jvi.01853-18>.
- Samadi, S., Wouters, I.M., Heederik, D.J.J., 2013. A review of bio-aerosol exposures and associated health effects in veterinary practice. *Ann. Agric. Environ. Med.* 20, 206–221.
- Schijven, J., Teunis, P., Suylen, T., Ketelaars, H., Hornstra, L., Rutjes, S., 2019. QMRA of adenovirus in drinking water at a drinking water treatment plant using UV and chlorine dioxide disinfection. *Water Res.* 158, 34–45. <https://doi.org/10.1016/j.watres.2019.03.090>.
- Schoen, M.E., Jahne, M.A., Garland, J.L., 2018. Human health impact of cross-connections in non-potable reuse systems. *Water* 10. <https://doi.org/10.3390/w10101352>.
- Scoizec, A., Niqueux, E., Thomas, R., Daniel, P., Schmitz, A., Le Bouquin, S., 2018. Airborne detection of H5N8 highly pathogenic avian influenza virus genome in poultry farms, France. *Front. Vet. Sci.* 5 <https://doi.org/10.3389/fvets.2018.00015>.
- Shi, K.W., Wang, C.W., Jiang, S.C., 2018. Quantitative microbial risk assessment of Greywater on-site reuse. *Sci. Total Environ.* 635, 1507–1519. <https://doi.org/10.1016/j.scitotenv.2018.04.197>.
- Sommer, J., Bobal, M., Bromberger, B., Mester, P.J., Rossmann, P., 2021. A new long-term sampling approach to viruses on surfaces. *Sci. Rep.* 11, 1–7. <https://doi.org/10.1038/s41598-021-96873-9>.
- Stobnicka-Kupiec, A., Gołofit-Szymczak, M., Cyprowski, M., Górny, R.L., 2022. Detection and identification of potentially infectious gastrointestinal and respiratory viruses at workplaces of wastewater treatment plants with viability qPCR/RT-qPCR. *Sci. Rep.* 12, 1–16. <https://doi.org/10.1038/s41598-022-08452-1>.
- Strubbia, S., Schaeffer, J., Besnard, A., Wacrenier, C., Le Menec, C., Garry, P., Desdouts, M., Le Guyader, F.S., 2020. Metagenomic to evaluate norovirus genomic diversity in oysters: impact on hexamer selection and targeted capture-based enrichment. *Int. J. Food Microbiol.* 323, 108588 <https://doi.org/10.1016/j.ijfoodmicro.2020.108588>.
- Tanner, B.D., Brooks, J.P., Gerba, C.P., Haas, C.N., Josephson, K.L., Pepper, I.L., 2008. Estimated occupational risk from bioaerosols generated during land application of Class B biosolids. *J. Environ. Qual.* 37, 2311–2321. <https://doi.org/10.2134/jeq2007.0193>.
- Teunis, P., Schijven, J., Rutjes, S., 2016. A generalized dose-response relationship for adenovirus infection and illness by exposure pathway. *Epidemiol. Infect.* 144, 3461–3473. <https://doi.org/10.1017/S0950268816001862>.
- Thorn, J., Kerekes, E., 2001. Health effects among employees in sewage treatment plants: a literature survey. *Am. J. Ind. Med.* 40, 170–179. <https://doi.org/10.1002/ajim.1085>.
- U.S. EPA, 2005. Occurrence and Exposure Assessment for the Final Long Term 2 Enhanced Surface Water Treatment Rule. EPA 815-R-06-002 EPA 815-R-06-002.
- Weesendorp, E., Stegeman, A., Loeffen, W.L.A., 2009. Quantification of classical swine fever virus in aerosols originating from pigs infected with strains of high, moderate or low virulence. *Vet. Microbiol.* 135, 222–230. <https://doi.org/10.1016/j.vetmic.2008.09.073>.
- WHO, 2008. Guidelines for Drinking-Water Quality: Incorporating 1st and 2nd Addenda, vol. 1. World Health Organization.
- Wilson, A.M., Verhoughstraete, M.P., Beamer, P.I., King, M.F., Reynolds, K.A., Gerba, C.P., 2021. Frequency of hand-to-head, -mouth, -eyes, and -nose contacts for adults and children during eating and non-eating macro-activities. *J. Expo. Sci. Environ. Epidemiol.* 31, 34–44. <https://doi.org/10.1038/s41370-020-0249-8>.
- Yan, C., Gui, Z., Cheng, W., Wu, J., ting, 2021. Quantitative microbial risk assessment of bioaerosols in a wastewater treatment plant by using two aeration modes. *Environ. Sci. Pollut. Res.* 28, 8140–8150. <https://doi.org/10.1007/s11356-020-11180-6>.
- Zaneti, R.N., Girardi, V., Spilki, F.R., Mena, K., Westphalen, A.P.C., da Costa Colares, E. R., Pozzebon, A.G., Etchepare, R.G., 2021. Quantitative microbial risk assessment of SARS-CoV-2 for workers in wastewater treatment plants. *Sci. Total Environ.* 754, 142163 <https://doi.org/10.1016/j.scitotenv.2020.142163>.

Discussió



L'aigua residual: font de contaminació i eina per a estudis epidemiològics. Detecció i caracterització de coronavirus en un context de pandèmia.

L'epidemiologia basada en aigües residuals ha demostrat ser una eina eficaç per al monitoratge de virus que circulen en una població. Amb la recent pandèmia de la COVID-19, aquesta eina de vigilància ha contribuït de forma significativa al seguiment de l'evolució epidemiològica. A més, les tècniques de NGS permeten estudiar el viroma de les aigües residuals i, aplicant l'assaig adequat, és possible obtenir la diversitat de variants, en aquest cas de SARS-CoV-2, que circulen a la població en un moment determinat.

En el primer article presentat en aquesta tesi consisteix es van revisar i avaluar les diferents metodologies utilitzades en mostres d'aigua residual per estudiar la disseminació de les variants del SARS-CoV-2. L'aparició de les variants α , β , γ i δ , considerades VOC en el seu moment, va evidenciar la necessitat de desenvolupar estratègies per estudiar-ne la circulació. Aquestes variants, associades amb les fluctuacions observades amb les onades de la pandèmia, presentaven mutacions signatura característiques, la majoria localitzades al gen de l'espícula vírica, que influïen en la infectivitat vírica i l'antigenicitat (Harvey et al., 2021). Per tant, les regions amb mutacions signatura de les VOC i VOI són generalment seleccionades com a diana per a les diferents estratègies analítiques.

A més, en aquest article es van incloure dades preliminars de dues aproximacions desenvolupades: assaigs específics de RT-nPCR amb posterior seqüenciació Sanger dissenyats per identificar mutacions signatura de les principals VOC i VOI, descrites fins a la data de la publicació de l'article, i una estratègia de seqüenciació massiva d'amplicons de tres regions diferents del gen de l'espícula vírica. Les dues

aproximacions es van aplicar de forma paral·lela en mostres compostes d'aigua residual recollides d'EDARs catalanes de diferents dimensions durant els mesos de febrer i maig de 2021.

Els assaigs específics nPCR i la posterior seqüenciació Sanger van permetre la identificació de mutacions signatura predominants en les mostres d'aigua residual, incloent-hi Del69/70, Del144, K417N i E484K. En canvi, mitjançant la seqüenciació massiva d'amplicons, es va obtenir la diversitat genètica de cada mostra, revelant diferents combinacions de mutacions signatura compatibles amb diferents variants, tal com s'espera de la diversitat vírica present en les aigües residuals. Curiosament, la seqüenciació massiva d'amplicons va permetre identificar el moment en què la variant α probablement es va convertir en la predominant a Catalunya, entre la primera i la segona setmana de febrer de l'any 2021. Aquests resultats van ser coherents amb els de la seqüenciació Sanger, ja que les mutacions signatura de la variant α només es van detectar quan les mostres presentaven un alt percentatge d'aquestes mutacions, és a dir, quan aquestes mutacions eren les predominants entre la barreja de seqüències corresponents a les diferents variants circulants.

Amb aquests resultats preliminars i la revisió de les diferents metodologies utilitzades per caracteritzar les variants de SARS-CoV-2 en aigües residuals, es va poder realitzar una avaluació assenyalant els avantatges i inconvenients de cadascuna. Els assaigs de RT-qPCR i RT-ddPCR dissenyats per detectar mutacions signatura d'una variant concreta són els més ràpids i permeten monitoritzar una variant en una regió on s'ha propagat i s'ha establert. L'estratègia RT-nPCR amb posterior seqüenciació Sanger pot ser útil per elucidar la variant predominant que circula en una regió, ja que amb aquest tipus de seqüenciació només s'obté la seqüència amplificada més abundant. En canvi, considerant que les aigües residuals contenen una barreja de variants, les tècniques de NGS són les més adequades si es vol realitzar una caracterització de la

diversitat present, perquè permeten identificar la diversitat de variants de SARS-CoV-2 que circulen en una regió determinada en un moment concret. En conclusió, cadascuna de les metodologies descrites ofereix diferents tipus d'informació, amb els seus propis avantatges i inconvenients. És important considerar aquests aspectes a l'hora de seleccionar l'estratègia més adequada, la qual dependrà dels objectius específics de l'estudi.

El segon estudi es va centrar en l'aplicació de tècniques de NGS per descriure l'ocurrència del SARS-CoV-2, juntament amb altres virus excretats que constitueixen el viroma de les aigües residuals. El SARS-CoV-2 s'excreta en femta per la majoria de les persones infectades, tinguin símptomes o no, de manera que el seu material genètic pot detectar fins a dies i setmanes després (Ahmed et al., 2020b; Medema et al., 2020). En aquest context, per tant, el SARS-CoV-2 podria considerar-se un membre més del viroma de les aigües residuals. No obstant, tenint en compte que l'aigua residual és una matriu complexa amb una gran diversitat vírica (Fernandez-Cassi et al., 2018; Martínez-Puchol et al., 2020a), la detecció de SARS-CoV-2 en estudis de viroma pot suposar un repte.

Es van estudiar dues mostres compostes de 24 hores d'aigua residual d'una EDAR de Barcelona. Una mostra va ser obtinguda al març de 2020, quan es va produir el pic de casos de la primera onada de la pandèmia a Catalunya, mentre que l'altra corresponia al moment abans de la segona onada, al juliol del 2020. La caracterització del viroma d'aquestes dues mostres es va fer tant amb com sense enriquiment de dianes, utilitzant el panell VirCapSeq-VERT per a la captura de seqüències de virus de vertebrats.

Les seqüències víriques de la família *Coronaviridae* només es van detectar en la mostra de març mitjançant la seqüenciació amb enriquiment

de dianes. Es van identificar coronavirus humans, incloent-hi SARS-CoV-2 i HCoV-OC43, associat amb el refredat comú (Van Der Hoek, 2007), així com coronavirus animals que infecten animals domèstics i rosegadors. Les seqüències de SARS-CoV-2 es van comparar amb les seqüències obtingudes de mostres clíniques que es van analitzar amb el mateix protocol de captura. A part de tres diferències nucleotídiques, els resultats obtinguts van demostrar una alta similitud entre les seqüències ambientals i clíniques. A causa del baix *coverage* obtingut del genoma, no es va poder determinar si aquestes diferències nucleotídiques estaven realment presents a l'ambient o si es tractaven d'errors de la pròpia seqüenciació. En aquest cas, es demostra que per a la caracterització de la diversitat de variants de SARS-CoV-2 circulant, pot ser més convenient la seqüenciació massiva d'amplicons, estratègia aplicada i avaluada a l'article anterior.

No es van detectar seqüències de coronavirus a la mostra de juliol, la qual presentava quantificacions més baixes de SARS-CoV-2 mitjançant qPCR, en consonància amb la baixa incidència reportada de la població en aquell moment de la pandèmia. Algunes possibles explicacions per no haver detectat altres coronavirus humans o animals podrien ser l'estacionalitat d'aquests virus i/o l'ús massiu de mascaretes, que implicarien una menor circulació vírica entre la població.

La seqüenciació massiva amb captura mitjançant el panell VirCapSeq-VERT ha permès identificar coronavirus humans i animals presents en aigua residual. Considerant que els coronavirus són patògens zoonòtics amb capacitat de causar brots, tant en humans com en animals, mitjançant la transmissió entre espècies, es requereix adoptar una perspectiva *One Health*, integrant la salut humana, animal i ambiental, per a la implementació d'estratègies efectives de control (Sharun et al., 2021). No obstant, el baix nombre de seqüències obtingudes suggereix que encara es requereixen millores per poder detectar aquests virus en

mostres ambientals. Una opció de millora és l'ús de panells dissenyats específicament per capturar membres de la família *Coronaviridae*, resultant en una major profunditat de seqüenciació i proporcionant una visió més completa de la diversitat de virus d'interès.

En el tercer i últim article d'aquest capítol es va desenvolupar i validar un panell de captura dirigit a millorar la seqüenciació de coronavirus en mostres ambientals i d'origen animal. El panell es va dissenyar amb sondes específiques per capturar diversos membres de la família *Coronaviridae* que infecten humans i animals, incloent-hi porcs, bestiar boví, rosegadors, gats, gossos, aus i ratpenats.

La seqüenciació massiva amb la captura i enriquiment de coronavirus utilitzant el panell dissenyat es va aplicar a tres mostres compostes de 24 hores d'aigua residual d'una EDAR de Barcelona, a quatre *pools* de mostres de guano de ratpenat recollides en ubicacions diferents de Catalunya i les Illes Balears i, finalment, a mostres de lixiviats de pollastres, porcs i conills procedents de camions de cadàvers animals.

En les mostres analitzades, gràcies al panell de captura dissenyat, es va poder seqüenciar una gran diversitat de coronavirus. Amb la seqüenciació HTS, sense enriquiment de dianes, només es van obtenir *reads* corresponents a coronavirus en un dels *pools* de mostres de guano, tot i que amb un ordre de magnitud molt inferior al nombre de *reads* obtinguts del mateix *pool* seqüenciat amb captura. Les mostres d'aigua residual van ser les que van mostrar la major diversitat de coronavirus, la qual cosa concorda amb el fet que aquest tipus mostres presenten una extensa diversitat vírica (Fernandez-Cassi et al., 2018; Martínez-Puchol et al., 2020a). Els coronavirus humans detectats, SARS-CoV-2, HCoV-229E i HCoV-OC43, poden atribuir-se a l'estat epidemiològic del moment, ja que les mostres van ser recollides durant un període en què la pandèmia causada pel SARS-CoV-2 continuava activa (de novembre de 2021 a

gener de 2022), mentre que el HCoV-229E i el HCoV-OC43 estan associats a l'ocurrència del refredat comú (Van Der Hoek, 2007). Es va detectar una diversitat més gran de coronavirus animals, essent els coronavirus de gos i de gat els que van presentar un major nombre de *reads* i que es podria explicar per l'eliminació de residus d'aquests animals domèstics a través del vàter o per la neteja de carrers.

Pel que fa a les mostres d'origen animal, la seqüenciació amb el panell de captura va permetre la identificació de coronavirus animals en les mostres analitzades. Es van detectar quatre espècies diferents de coronavirus que infecten ratpenats a les mostres de guano, així com el coronavirus de civeta HKU8, el qual s'ha especulat que podria estar associat a un possible esdeveniment de transmissió entre espècies a partir d'un coronavirus de ratpenat (He et al., 2022). Es van obtenir *reads* corresponents a coronavirus que infecten conills, aus, porcs, bestiar boví i gossos de les diferents mostres de lixiviats animals, tot i que amb una diversitat vírica i un *coverage* inferior que suggereix que aquestes matrius complexes tenen una ocurrència menor de coronavirus, un major grau de degradació dels àcids nucleics, o bé requereixen d'una metodologia optimitzada per poder-los detectar.

Així doncs, amb aquest últim article, l'aplicació del panell de captura dirigit a la seqüenciació de coronavirus ha representat una millora considerable en comparació amb la seqüenciació de mostres d'aigua residual utilitzant el panell comercial utilitzat a l'article anterior, VirCapSeq-VERT, dirigit a la captura de seqüències de virus de vertebrats. A més, també s'ha demostrat el seu potencial ús en mostres d'origen animal, concretament, guano i lixiviats animals. S'ha validat la seva eficàcia no només en la captura i seqüenciació de les espècies víriques incloses al disseny del panell, sinó també d'altres espècies dins de la família. Aquest panell de captura permet obtenir una major profunditat de seqüenciació per a la detecció de coronavirus d'interès i podria facilitar la descoberta de virus

nous, representant una eina eficaç per a l'estudi de virus zoonòtics i la identificació de noves espècies susceptibles a la transmissió entre espècies. L'estudi de les interaccions entre humans, animals i l'ambient compartit és essencial per identificar els potencials reservoris i hostes intermediaris de coronavirus, la qual cosa contribueix a la prevenció d'esdeveniments zoonòtics i malalties infeccioses emergents mitjançant la detecció primerenca de brots potencials i el desenvolupament de mesures eficaces (Islam et al., 2022). La seqüenciació massiva amb el panell de captura desenvolupat ofereix una millor comprensió de la dinàmica de les poblacions víriques i les seves possibles implicacions en la transmissió zoonòtica.

Exposició a virus contaminants d'aigües i aliments: detecció i caracterització vírica.

Els virus excretats a les aigües residuals poden ser resistents als tractaments de les EDARs i acabar contaminant les fonts d'aigua que s'utilitzen per beure, per regar o per a ús recreatiu, suposant així potencials vies de transmissió d'aquests virus. Aquest capítol de la tesi es centra en estudiar l'exposició a virus patògens que contaminen aigües i aliments.

Les fruites i verdures són susceptibles a la contaminació microbiològica en cada etapa de la producció alimentària. Amb l'increment del consum d'aquest tipus d'aliments en els últims anys, sovint consumits crus o mínimament processats, s'ha associat un augment en brots d'infeccions i malalties d'origen alimentari, la majoria dels quals s'han vinculat a virus (Balali et al., 2020; Callejón et al., 2015; Fernandez-Cassi et al., 2017). L'aigua utilitzada per regar aquests aliments és un factor crític, ja que constitueix una font potencial de virus patògens transmesos pels aliments (Garcia et al., 2015; Shaheen et al., 2019). Per aquest motiu, en el quart

article es van estudiar mostres d'enciam, julivert i maduixes procedents de tres llocs diferents de producció d'agricultura ecològica situats a la província de Barcelona, juntament amb les seves corresponents aigües de reg, incloent-hi mostres d'aigua subterrània (llocs de producció 1 i 2) i d'aigua de riu (lloc de producció 3), així com d'altres fonts alternatives com el riu Llobregat o l'aigua de tractament secundari i de terciari. Aquestes últimes es van incloure perquè l'aigua regenerada està esdevenint una font important d'aigua d'irrigació i s'està potenciant el seu ús en contextos d'escassetat d'aigua. Totes les mostres es van analitzar mitjançant assaigs qPCR específics de l'indicador de contaminació fecal HAdV i altres patògens rellevants, com ara NoV, per tal de detectar i quantificar el nivell de contaminació vírica. Les mostres amb una major contaminació vírica es van seleccionar per ser analitzades mitjançant dues tècniques de NGS: la seqüenciació massiva amb enriquiment de dianes, utilitzant VirCapSeq-VERT com a panell de captura, i la seqüenciació massiva d'amplicons dirigida a la detecció i caracterització de NoV, considerat una de les principals causes dels brots d'origen alimentari (Gobeil et al., 2020; Gonzales-Gustavson et al., 2019; Koo et al., 2010), i de HPV, el qual s'ha descrit la seva ocurrència en aigües residuals i fluvials (La Rosa et al., 2013; Martínez-Puchol et al., 2020a, 2021; Rusiñol et al., 2020), però la seva transmissió a través de l'aigua contaminada encara és incerta.

Els assaigs de qPCR van revelar una contaminació baixa però freqüent de patògens vírics HAdV i/o NoV en el 46.9% de les mostres d'aliments (6/12 mostres d'enciam, 4/12 mostres de maduixa i 5/8 mostres de julivert) i en el 50% de les mostres d'aigües de reg analitzades.

La seqüenciació massiva amb captura utilitzant el panell VirCapSeq-VERT va permetre la detecció de virus humans i animals en totes les mostres d'aigua de reg analitzades. La majoria de *reads* obtinguts van correspondre a virus humans en totes aquestes mostres, suposant inclús la totalitat, o gairebé la totalitat, en les mostres d'aigua de tractament

secundari i de terciari, respectivament. En la resta de fonts d'aigües de reg, es va detectar una proporció major de *reads* corresponents a virus d'animals de diferents famílies víriques, incloent-hi *Astroviridae*, *Caliciviridae*, *Genomoviridae* i *Parvoviridae*, d'entre altres, amb una gran diversitat d'hostes als que infecten, com ara porcs, animals domèstics, cavalls, bestiar boví, aus i ratpenats. L'origen d'aquests virus podria estar relacionat amb la filtració d'aigües residuals o la contaminació provinent de zones de producció de bestiar properes als llocs d'estudi (Wei et al., 2011). En les mostres d'aliments, es va obtenir un nombre de *reads* molt baix de virus de vertebrats només en una mostra de julivert, els quals es van identificar com a NoV GII.17, un genotip que va emergir com a causa principal de brots de gastroenteritis a la Xina i al Japó durant l'hivern del 2014/2015 (de Graaf et al., 2015). No es van detectar virus humans en altres mostres d'aliments mitjançant la seqüenciació amb VirCapSeq-VERT, tot i haver detectat contaminació vírica mitjançant qPCR. Aquest fet podria explicar-se per la complexitat d'aquestes mostres que dificulta la detecció de virus, tal i com s'ha descrit anteriorment en el cas de fruits de tipus baia, per presència d'inhibidors i un baix pH (Le Guyader et al., 2004), o també per limitacions associades amb el protocol SISPA (Duhaime et al., 2012; Fernandez-Cassi et al., 2017; Karlsson et al., 2013). Per tal de millorar la detecció de virus en aquestes matrius alimentàries es podria optimitzar el processament de les mostres per eliminar possibles inhibidors, dissenyar panells de captura dirigits a la seqüenciació massiva de virus patògens transmesos per aigua i aliments o, alternativament, utilitzar l'estratègia de seqüenciació massiva d'amplicons que permetés la caracterització vírica dins d'una mateixa grup de virus, detallada a continuació.

La seqüenciació massiva d'amplicons va permetre l'estudi de la diversitat vírica de NoV i HPV en les mostres d'aliments i d'aigües de reg analitzades. Les mostres d'aigua residual tractada i les de riu són les que

van presentar una major diversitat d'aquests virus. En general, totes les mostres d'aliments van presentar una menor diversitat de NoV i HPV, però és necessari destacar que tots els tipus vírics identificats en aliments també es van trobar en les fonts d'aigua de reg. El genotip NoV GII.4, considerat actualment el genotip més abundant i el principal causant dels brots de norovirus a nivell global (CDC, 2023), es va identificar en totes les fonts d'aigua de reg, així com en mostres de maduixa i de julivert. La majoria dels HPV detectats pertanyen al gènere Betapapillomavirus, i cal destacar el HPV-8 seqüenciat en les mostres d'aigua tractada, el qual s'ha associat amb el carcinoma de cèl·lules escamoses (SCC) (Li et al., 2013).

Els resultats d'aquest estudi indiquen que els aliments de producció orgànica i les aigües d'irrigació presenten una àmplia diversitat de virus patògens que podrien representar un risc per als humans. A més, els resultats assenyalen que l'aigua de reg és la font més probable dels virus patògens detectats en els aliments analitzats, especialment de NoV i HPV. La detecció freqüent de HPV tant en aliments com en les aigües de reg suggereix que podria considerar-se un potencial patògen emergent, i s'hauria d'estudiar més exhaustivament la seva transmissió a través de les aigües de reg i els aliments. La seqüenciació massiva utilitzant el panell de captura VirCapSeq-VERT ha demostrat ser eficaç per a la detecció de virus humans i animals que integren el viroma de les aigües de reg, mentre que la seqüenciació massiva d'amplicons ha permès caracteritzar els patògens vírics NoV i HPV que contaminen tant les aigües de reg com els aliments de producció orgànica. Ambdues tècniques de NGS han permès detectar virus patògens humans fins i tot en mostres amb baixos nivells de contaminació vírica detectats mitjançant qPCR, indicant que aquestes tècniques podrien ser útils per a identificar i caracteritzar patògens vírics en el context de la salut alimentària. D'aquesta manera, són eines potencials per monitoritzar virus patògens

que poden contaminar els aliments i les aigües de reg, amb l'objectiu de prevenir i reduir els brots d'infeccions i de malalties d'origen alimentari.

Seguint amb l'exposició a virus que contaminen les fonts d'aigua, el cinquè article va avaluar la contaminació vírica i va explorar la diversitat de virus presents en aigües subterrànies i d'escorrentia d'un entorn urbà. L'ús dels aqüífers urbans com a fonts d'aigua potable està augmentant anualment, especialment durant els períodes de sequera, però la monitorització de la qualitat de les aigües subterrànies és insuficient, en molts casos, en quant a controls d'elements biològics (Idescat, 2022). Aquests aqüífers urbans són susceptibles a la contaminació per patògens a través de diverses fonts de contaminació, com ara les fuites de clavegueram o les escorrenties urbanes (Rusiñol, 2023).

En aquest estudi, es va analitzar la presència de virus indicadors de contaminació fecal d'origen humà i patògens vírics en aigües residuals i aigües subterrànies urbanes de quatre localitzacions de Barcelona amb diferent grau d'urbanització, així com en aigües d'escorrenties mitjançant assaigs de qPCR. A més, es va aplicar la seqüenciació massiva amb enriquiment de dianes mitjançant el panell de captura VirCapSeq-VERT per explorar la diversitat de virus de vertebrats present en les mostres d'aigua subterrània i d'escorrentia, oferint informació sobre l'origen de la contaminació vírica i les possibles rutes de transmissió de virus en àrees urbanes.

Es va identificar HAdV en totes les mostres d'aigua residual, en el 67% de les mostres d'aigua subterrània i en una mostra d'escorrentia mitjançant qPCR, indicant contaminació fecal humana. El patògen víric NoV GI es va detectar en aigües residuals i en dues mostres d'aigua subterrània d'hivern de zones altament i mitjanament urbanitzades.

La seqüenciació massiva utilitzant el panell VirCapSeq-VERT va permetre la detecció de virus humans i animals en aigües subterrànies i

d'escorrentia. En les aigües subterrànies, es va identificar el HPV-38, relacionat amb càncer de pell (Kocjan et al., 2009), així com circovirus i picornavirus animals que infecten ratpenats, coloms o gossos. D'altra banda, les mostres d'escorrentia van mostrar una major diversitat vírica, incloent-hi virus que infecten humans, coloms, animals domèstics o ratpenats, i que pertanyen a les famílies víriques *Anelloviridae*, *Astroviridae*, *Caliciviridae*, *Circoviridae*, *Coronaviridae*, *Genomoviridae*, *Parvoviridae*, *Picobirnaviridae*, *Picornaviridae*, *Poxviridae* i *Retroviridae*. Altres estudis havien descrit anteriorment la presència de virus en aquest tipus de matriu ambiental mitjançant assaigs de PCR o qPCR (Sauer et al., 2011; Sidhu et al., 2012, 2013). Aquests estudis van revelar un perfil de virus de transmissió fecal-oral compatible amb la contaminació provinent de les aigües residuals durant episodis de pluja i, per tant, suggerint una exposició potencial a una àmplia varietat de virus excretats.

Els resultats d'aquest estudi aporten informació sobre el viroma present en aigües subterrànies urbanes i algunes de les seves fonts de contaminació, com són les aigües residuals i les escorrenties urbanes. Amb la creixent demanda dels recursos hídrics i l'ampliació dels usos de les aigües subterrànies urbanes, és important identificar els virus patògens que poden contaminar-les, considerat un camp encara emergent i amb estudis limitats (Rusiñol, 2023). La seqüenciació massiva amb enriquiment de dianes ha permès detectar virus humans i animals en aigües subterrànies i d'escorrentia urbanes, contribuint a avaluar els riscos potencials associats amb les malalties transmeses per l'aigua, especialment en àrees urbanes cada vegada més poblades, però també en àrees rurals que depenen de l'ús d'aigües subterrànies per a beure o per al reg. Monitoritzar els virus humans i animals amb potencial zoonòtic en aquest tipus d'aigües és essencial per identificar les fonts de contaminació i, així, poder desenvolupar mesures de mitigació efectives.

Identificació de virus presents en entorns de treball amb risc d'exposició. Avaluació del risc ocupacional.

Considerant que l'exposició ocupacional a patògens pot comportar riscos per a la salut dels treballadors, en el sisè article es va avaluar l'exposició vírica en dos entorns de treball amb risc d'exposició.

La majoria de les condicions de salut adverses reportades pels treballadors d'EDARs s'han relacionat amb l'exposició a bioaerosols (Amoah et al., 2022; Carducci et al., 2016; Corrao et al., 2012). Aquests bioaerosols es poden generar a través de diferents processos durant el tractament de les aigües residuals, de manera que microorganismes patògens de les aigües residuals poden ser alliberats fàcilment a l'atmosfera (Kataki et al., 2022) i, posteriorment, dipositar-se sobre les superfícies (Han et al., 2013).

Amb la ràpida expansió de la producció porcina intensiva i a gran escala, l'emissió d'aerosols en les granges de porcs també s'ha convertit en una preocupació creixent (Liu et al., 2023). S'han detectat virus patògens a l'aire d'aquestes granges, incloent-hi PRRSV, PEDV, CSFV, ASFV i IAV (Alonso et al., 2015), aquest últim també identificat en superfícies (Neira et al., 2016).

Així doncs, tant els treballadors d'EDARs com els de granges de porcs poden estar exposats a virus, ja sigui per la inhalació de bioaerosols o per la ingestió oral després del contacte directe amb superfícies, roba o eines contaminades. Per aquest motiu, es van analitzar mostres d'aire i de superfícies d'aquests entorns ocupacionals a Catalunya per tal d'identificar els virus als quals els treballadors estan exposats. Les mostres d'aerosols es van obtenir amb el mostrejador d'aire ciclònic Coriolis μ , mentre que les superfícies es van mostrejar seguint l'estratègia descrita per Sommer i col·laboradors per estudiar virus en superfícies a llarg termini (Sommer et al., 2021), utilitzant adhesius de paper en

superfícies tocadess freqüentment pels treballadors. Les campanyes de mostreigs van incloure hivern i estiu. En el cas de l'EDAR, les mostres d'aire es van recollir tant en zones interiors com en exteriors, mentre que per al mostreig de superfícies es van distingir dos escenaris diferents: per una banda, els espais de treball, i per l'altra, la sala que els treballadors utilitzen com a cuina durant el seu temps de descans. Totes les mostres es van analitzar mitjançant assaigs de qPCR per tal d'avaluar la contaminació vírica, aspecte que es va tenir en compte per fer el *pools* de mostres per a la seqüenciació massiva de virus de vertebrats amb el panell de captura VirCapSeq-VERT.

Els HAdV i PAdV es van detectar en un gran percentatge de mostres d'aerosols i de superfícies de l'EDAR i de la granja de porcs, respectivament, mitjançant qPCR. Són virus persistents que s'excreten i es troben freqüentment a l'ambient (Hijnen et al., 2006), els HAdV es consideren indicadors de contaminació fecal humana (Bofill-Mas et al., 2013), mentre que els PAdV s'han utilitzat per identificar fonts de contaminació fecal porcina (Rusiñol et al., 2014). Totes les mostres d'aire i de superfícies van resultar negatives per a l'assaig de SARS-CoV-2, excepte dues mostres de superfícies de la granja de porcs. La detecció de SARS-CoV-2 en aquestes mostres es va relacionar amb un brot existent entre els treballadors durant el període de mostreig.

La seqüenciació massiva amb el panell de captura VirCapSeq-VERT va permetre la detecció de virus humans i animals en les mostres de l'EDAR. El perfil de virus detectats en les mostres d'aerosols correspon a les famílies víriques descrites habitualment en les aigües residuals, suggerint que la font d'aquests virus és l'aerosolització de les aigües residuals. Alguns exemples d'aquestes famílies són *Adenoviridae*, *Astroviridae*, *Caliciviridae*, *Circoviridae* i *Parvoviridae*, les quals es van detectar tant en mostres d'aerosols d'interior com d'exterior. De fet, es va identificar el HAdV-41 com el principal serotip d'HAdV detectat en totes les mostres

d'aerosols i de superfícies, coincidint amb el serotip obtingut per seqüenciació Sanger de mostres d'aigua residual de l'EDAR durant el mateix període de mostreig i, per tant, reforçant la idea que l'aerosolització de l'aigua residual és la font dels virus patògens detectats. L'HAdV-41 està associat etiològicament a gastroenteritis, i la seva alta prevalença a l'ambient s'ha identificat anteriorment (Bofill-Mas et al., 2013). La resta de virus humans detectats es relacionen amb infeccions asimptomàtiques o amb un ampli rang de símptomes, com ara gastroenteritis, hepatitis i meningitis (Rusiñol & Girones, 2017). En aerosols d'exterior es van identificar seqüències corresponents a HPV, considerats epiteliotròpics històricament, però la seva ocurrència en aigües residuals (Iaconelli et al., 2015; La Rosa et al., 2013; Martínez-Puchol et al., 2020a; Rusiñol et al., 2020) i la seva excreció en femta (Di Bonito et al., 2015; Iaconelli et al., 2015; La Rosa et al., 2013), juntament amb la diversitat vírica detectada en aliments i aigües de reg en el quart article d'aquesta tesi, suggereixen una nova via de transmissió fecal-oral que cal investigar amb més profunditat. Per últim, la diversitat de virus que infecten aus identificada tant en mostres d'aerosols com en mostres de superfícies concorda amb resultats previs que van descriure virus avians en mostres d'aigua residual de la mateixa EDAR i confirma l'impacte dels efluents de la indústria avícola (Carratalà et al., 2012).

En canvi, en les mostres de superfícies, a més de detectar alguns virus associats a l'aigua, es va identificar un perfil de virus que podria ser compatible amb el contacte de les mans contaminades dels treballadors. Aquest fet s'evidencia amb la gran diversitat de HPV identificada en aquestes mostres, o amb la detecció del virus d'Epstein-Barr (EBV) i el poliomavirus de les cèl·lules de Merkel (MCPyV). EBV i MCPyV són virus altament prevalents a la població que causen infeccions asimptomàtiques, però en alguns casos poden causar lesions cutànies i càncer en humans

(Huang et al., 2023; Liu et al., 2016; Sanguenza-Acosta & Sandoval-Romero, 2018).

La seqüenciació massiva amb el panell de captura també va possibilitar la detecció de virus de vertebrats en les mostres d'aerosols i superfícies de la granja de porcs, la majoria essent virus comuns que infecten porcs i que pertanyen a famílies com *Adenoviridae*, *Astroviridae*, *Circoviridae*, *Herpesviridae*, *Parvoviridae* i *Picornaviridae*. Per tant, aquests resultats evidencien que els treballadors de granges de porcs estan exposats a una gran diversitat de virus animals. De fet, s'ha descrit que els treballadors d'aquests entorns tenen un major risc d'infeccions per virus de la grip zoonòtics (Gray et al., 2007) i, recentment, s'ha declarat un cas de grip porcina en humans a Catalunya (gener 2024), corresponent a una persona que treballava en una granja porcina (ASPCAT, 2024).

Les dades de les quantificacions de HAdV i PAdV en mostres d'aerosols i superfícies de l'EDAR i la granja de porcs, respectivament, obtingudes amb assaigs qPCR, es van utilitzar per estimar el risc d'aquests treballadors associat a l'exposició vírica mitjançant QMRA, considerant dues vies de transmissió: la inhalació d'aerosols i la ingestió oral a través de superfícies contaminades i el contacte mà-boca. En el cas de l'EDAR, es va seleccionar HAdV com a virus de referència per estimar el risc ocupacional associat a les infeccions per virus patògens presents a l'aigua residual. En canvi, en el cas de la granja, es va estimar el risc, amb una aproximació de QMRA, que tindrien els treballadors de granja d'infectar-se amb un virus hipotètic amb un potencial zoonòtic assumit i característiques similars a les del PAdV, sota un context *One Health*.

L'anàlisi de QMRA va revelar que l'exposició ocupacional a l'EDAR comporta un notable risc d'infeccions víriques si no es prenen mesures de seguretat. La probabilitat mitjana diària de malaltia per la inhalació d'HAdV va resultar ser més alta durant l'estiu (5.4%), mentre que el major risc

d'ingestió oral d'HAdV a través de superfícies contaminades es va observar als espais de treball durant l'hivern (13%).

A la granja de porcs, la simulació de QMRA va suggerir un possible risc en cas d'exposició a un virus zoonòtic. La probabilitat mitjana de malaltia diària més alta es donaria per la ingestió oral, mitjançant el contacte amb superfícies contaminades, durant l'estiu (63.6%), mentre que la inhalació seria similar entre l'hivern i l'estiu (20.8% i 21.8%).

Tots aquests resultats indiquen considerables riscos potencials per als treballadors d'EDARs i de granges de porcs, coincidint amb estudis epidemiològics que informen sobre símptomes relacionats amb aquests entorns ocupacionals i els efectes sobre la salut entre aquests treballadors (Al-Batanony & El-Shafie, 2011; Andersen et al., 2004; Donham et al., 1995; Douwes et al., 2001; Samadi et al., 2013; Thorn & Kerekes, 2001). De fet, les probabilitats mitjanes estacionals de malaltia en tots els escenaris estarien excedint el límit de referència proposat per l'EPA, 10^{-4} per persona i per any (U.S. EPA, 2005).

Les anàlisis de QMRA i NGS realitzades en aquest estudi han permès identificar l'exposició a virus patògens presents en entorns ocupacionals de risc i avaluar el risc associat a aquesta exposició. Aquesta informació és essencial per al desenvolupament i la implementació d'estratègies i mesures de seguretat per protegir els treballadors d'EDARs i de granges de porcs, com ara procediments de neteja efectius o l'ús d'EPI. La seqüenciació massiva amb enriquiment de dianes aplicada a mostres d'aerosols i superfícies d'aquests entorns ocupacionals pot ser una eina útil per detectar i caracteritzar els virus que circulen, així com identificar potencials virus emergents i zoonòtics. En aquest últim cas, especialment en mostres de granges de porcs, dissenyar un panell de captura dirigit a la seqüenciació de virus amb potencial zoonòtic i/o pandèmic podria millorar-ne la detecció.

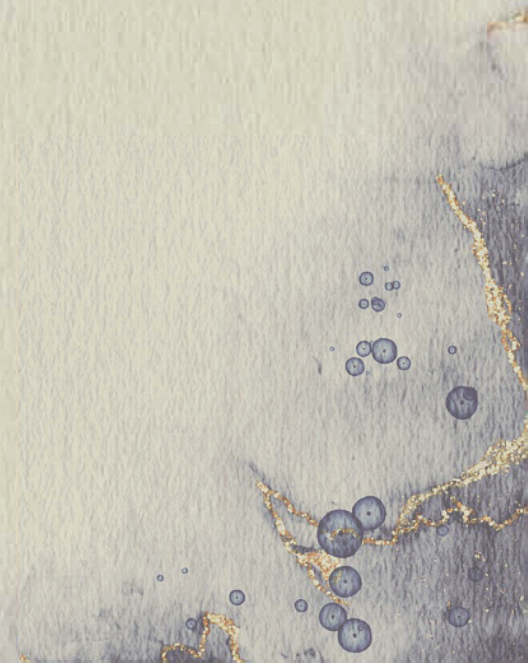
Els resultats d'aquesta tesi demostren que existeix una gran diversitat de virus presents en l'ambient als quals els humans estem exposats, constituint l'exposoma víric. És important estudiar l'ocurrència dels virus a l'ambient, les seves vies de transmissió, la seva epidemiologia i ecologia, així com preveure quins impactes pot tenir el canvi climàtic en tots aquests aspectes. S'ha demostrat que el canvi climàtic afecta els ecosistemes de tot el món, i constitueix un dels factors que impulsen l'aparició de virus emergents i zoonòtics. Un altre impacte directe són els episodis de sequera i d'escassetat d'aigua, en els quals s'ha de recórrer a la utilització de fonts alternatives d'aigua, com ara les aigües regenerades o les aigües subterrànies, les quals poden suposar una via de transmissió de virus patògens si no es monitoritzen per assegurar la seva qualitat.

En aquesta tesi no s'ha avaluat la capacitat infecciosa dels virus detectats, però els mètodes moleculars utilitzats han permès detectar el material genètic, l'empremta vírica a l'ambient i, per tant, el testimoni de la presència i existència d'aquests virus. Aquesta empremta molecular ens aporta informació important per identificar els virus circulants i les possibles fonts de contaminació, i així poder avaluar els potencials riscos associats. Tal i com s'ha discutit al llarg d'aquesta tesi, el monitoratge i la vigilància ambiental són clau per a una gestió eficaç dels brots actuals i futurs. La seqüenciació de virus d'interès en mostres ambientals pot suposar un repte, ja que sovint són mostres complexes amb una gran diversitat vírica que pot dificultar la detecció dels virus que es vulguin detectar, i que potser es troben en menor proporció a la resta. En aquesta tesi es demostra que la seqüenciació massiva amb enriquiment de dianes o la seqüenciació massiva d'amplicons són les aproximacions més efectives per abordar aquesta limitació. La seqüenciació amb enriquiment de dianes permet adaptar la captura en funció del panell utilitzat, podent escollir el grau d'especificitat, des d'utilitzar panells més generals, dirigits a virus de vertebrats, fins a panells més específics dissenyats per a la

captura de membres d'una família vírica concreta, com els coronavirus. D'altra banda, la seqüenciació massiva d'amplicons permet una caracterització vírica més sensible dins d'un grup determinat de virus.

El monitoratge ambiental a través d'aigües residuals i altres fonts de contaminació, aigües subterrànies, aliments de producció orgànica i aigües de reg, aerosols, superfícies i mostres d'origen animal, permet identificar els virus que circulen en una població determinada, contribuint així a la prevenció de malalties i al desenvolupament d'accions eficients, útils i econòmiques contra properes pandèmies. La interacció entre virus, ambient, animals i humans és complexa i, tot i que els virus són les entitats més abundants del planeta, encara queda molt per aprendre.

Conclusions

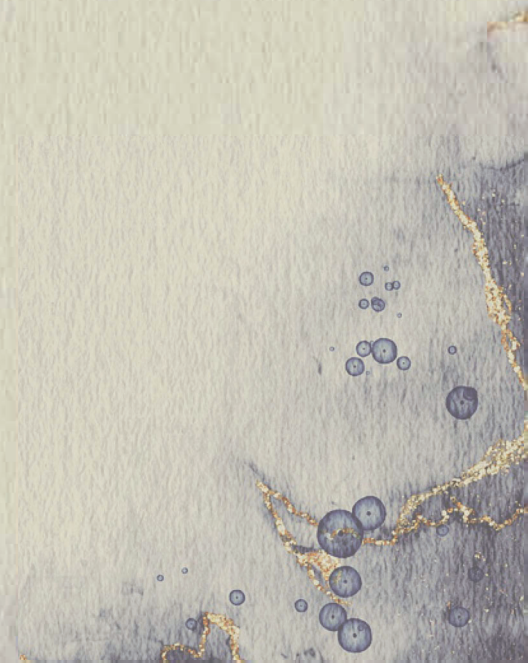


- L'aplicació de diferents tècniques de NGS ha permès la identificació i caracterització de virus en una gran varietat de matrius ambientals, incloent-hi aigües residuals, aigües d'escorrentia, aigües subterrànies, mostres d'origen animal, aliments de producció orgànica i aigües de reg, així com superfícies i aerosols d'entorns de treball. Els virus identificats constitueixen l'exposoma víric al qual els humans estem exposats.
- L'estudi del viroma de l'aigua residual en un context de pandèmia mitjançant la seqüenciació massiva amb VirCapSeq-VERT ha permès la detecció de coronavirus humans i animals, juntament amb altres virus que constitueixen aquest viroma. En canvi, l'ús d'un panell més específic dissenyat per capturar membres de la família *Coronaviridae* ha millorat la seqüenciació de coronavirus humans i animals en mostres d'aigua residual i ha demostrat el seu potencial ús en mostres d'origen animal.
- Les diferents metodologies per estudiar la diversitat de variants del SARS-CoV-2 en aigües residuals ofereixen diferents tipus d'informació, i la seva selecció s'ha de basar en els objectius específics de l'estudi.
- Els aliments de producció orgànica i les aigües d'irrigació mostren una àmplia diversitat de virus patògens que podrien representar un risc per als humans. A més, s'ha pogut determinar que l'aigua de reg és la font més probable dels virus patògens detectats en els aliments analitzats. La seqüenciació massiva amb VirCapSeq-VERT ha permès la detecció de virus humans i animals que integren el viroma de les aigües de reg, mentre que la seqüenciació massiva d'amplicons ha permès caracteritzar els patògens vírics NoV i HPV que contaminen tant les aigües de reg com els aliments de producció orgànica.

CONCLUSIONS

- L'aplicació de la seqüenciació massiva amb VirCapSeq-VERT ha permès estudiar el viroma de les aigües subterrànies i d'escorrentia urbanes, demostrant el seu potencial per identificar virus patògens i les seves possibles fonts de contaminació, així com per avaluar els riscos associats en un context d'escassetat d'aigua.
- La seqüenciació massiva amb VirCapSeq-VERT ha permès la identificació dels virus presents en l'aire i superfícies d'entorns ocupacionals de risc, contribuint a una millor comprensió dels riscos vírics als quals estan exposats els treballadors. S'ha pogut determinar que els treballadors d'EDARS es troben exposats a virus humans i animals provinents de l'aerosolització de l'aigua residual, mentre que els treballadors de granja estan exposats a una gran diversitat de virus animals.
- L'ús d'adhesius de paper, "gomets", ha demostrat ser una estratègia de mostreig eficient per a l'estudi de virus en superfícies, compatible amb posteriors anàlisis de qPCR, NGS i QMRA.
- L'avaluació quantitativa del risc microbiològic ha revelat que l'exposició ocupacional a l'EDAR comporta un notable risc d'infeccions víriques si no es prenen mesures de seguretat. La simulació de QMRA a la granja de porcs també suggereix un potencial risc en cas d'exposició a un virus zoonòtic hipotètic.
- El monitoratge i la vigilància ambiental per a la identificació de virus patògens, emergents i zoonòtics contribueix a una gestió eficaç dels brots actuals i futurs.

Referències



- Abdeldayem, O. M., Dabbish, A. M., Habashy, M. M., Mostafa, M. K., Elhefnawy, M., Amin, L., ... Rene, E. R. (2022). Viral outbreaks detection and surveillance using wastewater-based epidemiology, viral air sampling, and machine learning techniques: A comprehensive review and outlook. *Science of The Total Environment*, 803 (January), 149834. <https://doi.org/10.1016/j.scitotenv.2021.149834>
- Ahmed, A., Safdar, M., Sardar, S., Yousaf, S., Farooq, F., Raza, A., ... Afzal, S. (2022). Modern vaccine strategies for emerging zoonotic viruses. *Expert Review of Vaccines*, 21(12), 1711-1725. <https://doi.org/10.1080/14760584.2022.2148660>
- Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J. W., ... Mueller, J. F. (2020a). First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: A proof of concept for the wastewater surveillance of COVID-19 in the community. *Science of The Total Environment*, 728, 138764. <https://doi.org/10.1016/j.scitotenv.2020.138764>
- Ahmed, W., Bertsch, P. M., Bibby, K., Haramoto, E., Hewitt, J., Huygens, F., ... Bivins, A. (2020b). Decay of SARS-CoV-2 and surrogate murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-based epidemiology. *Environmental Research*, 191(August), 110092. <https://doi.org/10.1016/j.envres.2020.110092>
- Ahmed, W., Simpson, S. L., Bertsch, P. M., Bibby, K., Bivins, A., Blackall, L. L., ... Shanks, O. C. (2022). Minimizing errors in RT-PCR detection and quantification of SARS-CoV-2 RNA for wastewater surveillance. *Science of the Total Environment*, 805. <https://doi.org/10.1016/j.scitotenv.2021.149877>
- Al-Batanony, M. A., & El-Shafie, M. K. (2011). Work-related health effects among wastewater treatment plants workers. *International Journal of Occupational and Environmental Medicine*, 2(4), 237-244.
- Alonso, C., Raynor, P. C., Davies, P. R., & Torremorell, M. (2015). Concentration, size distribution, and infectivity of airborne particles carrying swine viruses. *PLoS ONE*, 10(8), 1-12. <https://doi.org/10.1371/journal.pone.0135675>

REFERÈNCIES

- Amoah, I. D., Kumari, S., & Bux, F. (2022). A probabilistic assessment of microbial infection risks due to occupational exposure to wastewater in a conventional activated sludge wastewater treatment plant. *Science of the Total Environment*, 843(June), 156849. <https://doi.org/10.1016/j.scitotenv.2022.156849>
- Anand, U., Bianco, F., Suresh, S., Tripathi, V., Núñez-Delgado, A., & Race, M. (2021). SARS-CoV-2 and other viruses in soil: An environmental outlook. *Environmental Research*, 198, 111297. <https://doi.org/10.1016/j.envres.2021.111297>
- Andersen, C. I., Von Essen, S. G., Smith, L. M., Spencer, J., Jolie, R., & Donham, K. J. (2004). Respiratory symptoms and airway obstruction in swine veterinarians: A persistent problem. *American Journal of Industrial Medicine*, 46(4), 386-392. <https://doi.org/10.1002/ajim.20080>
- Antwi-Agyei, P., Biran, A., Peasey, A., Bruce, J., & Ensink, J. (2016). A faecal exposure assessment of farm workers in Accra, Ghana: A cross sectional study. *BMC Public Health*, 16(1), 1-13. <https://doi.org/10.1186/s12889-016-3266-8>
- ARTIC Network (2020). *Real-time molecular epidemiology for outbreak response* [en línia]. Disponible a: <https://artic.network> [Últim accés: 09/04/2024].
- ASPCAT, Agència de Salut Pública de Catalunya (2022). *Xarxa de vigilància del SARS-CoV-2 en aigües residuals i les seves variants* [en línia]. Disponible a: <https://salutpublica.gencat.cat/ca/detalls/Article/Xarxa-de-monitoratge-de-la-presencia-del-SARS-CoV-2-en-aigües-residuals> [Últim accés: 09/04/2024].
- ASPCAT, Agència de Salut Pública de Catalunya (2024). *Nota informativa sobre la declaració d'un cas de grip porcina en humans a Catalunya (gener de 2024)* [en línia]. Disponible a: <https://salutpublica.gencat.cat/ca/detalls/Article/Nota-informativa-sobre-la-declaracio-dun-cas-en-huma-de-grip-porcina-a-Catalunya-gener-2024> [Últim accés: 09/04/2024]

- Aw, T. G., Wengert, S., & Rose, J. B. (2016). Metagenomic analysis of viruses associated with field-grown and retail lettuce identifies human and animal viruses. *International Journal of Food Microbiology*, 223, 50-56.
<https://doi.org/10.1016/j.ijfoodmicro.2016.02.008>
- Balali, G. I., Yar, D. D., Afua Dela, V. G., & Adjei-Kusi, P. (2020). Microbial Contamination, an Increasing Threat to the Consumption of Fresh Fruits and Vegetables in Today's World. *International Journal of Microbiology*, 2020.
<https://doi.org/10.1155/2020/3029295>
- Bar-Or, I., Weil, M., Indenbaum, V., Bucris, E., Bar-Ilan, D., Elul, M., ... Zuckerman, N. S. (2021). Detection of SARS-CoV-2 variants by genomic analysis of wastewater samples in Israel. *Science of the Total Environment*, 789, 148002. <https://doi.org/10.1016/j.scitotenv.2021.148002>
- Barbé, L., Schaeffer, J., Besnard, A., Jousse, S., Wurtzer, S., Moulin, L., ... Desdouits, M. (2022). SARS-CoV-2 Whole-Genome Sequencing Using Oxford Nanopore Technology for Variant Monitoring in Wastewaters. *Frontiers in Microbiology*, 13(June), 1-14.
<https://doi.org/10.3389/fmicb.2022.889811>
- Bardhan, M., Ray, I., Roy, S., Bhatt, P., Patel, S., Asri, S., ... Anand, A. (2023). Emerging zoonotic diseases and COVID-19 pandemic: global Perspective and Indian Scenario. *Annals of Medicine & Surgery*, 85(8), 3997-4004.
<https://doi.org/10.1097/ms9.0000000000001057>
- Bateman, A. (2019). UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Research*, 47(D1), D506-D515. <https://doi.org/10.1093/nar/gky1049>
- Biziagos, E., Passagot, J., Crance, J. M., & Deloince, R. (1988). Long-term survival of hepatitis A virus and poliovirus type 1 in mineral water. *Applied and Environmental Microbiology*, 54(11), 2705-2710.
<https://doi.org/10.1128/aem.54.11.2705-2710.1988>
- Blatchley, E. R., & Cui, H. (2023). Quantitative Microbial Risk Assessment for Quantification of the Effects of Ultraviolet Germicidal Irradiation on COVID-19 Transmission. *Environmental Science and Technology*, 57(45), 17393-17403. <https://doi.org/10.1021/acs.est.3c03026>

REFERÈNCIES

- Bobal, M., Witte, A. K., Mester, P., Fister, S., Schoder, D., & Rossmanith, P. (2019). A Novel Method for Sampling and Long-Term Monitoring of Microbes That Uses Stickers of Plain Paper. *Applied and environmental microbiology*, 85(14), 1-11. <https://doi.org/10.1128/AEM.00766-19>
- Bofill-Mas, S., Pina, S., & Girones, R. (2000). Documenting the epidemiologic patterns of polyomaviruses in human populations by studying their presence in urban sewage. *Applied and Environmental Microbiology*, 66(1), 238-245. <https://doi.org/10.1128/AEM.66.1.238-245.2000>
- Bofill-Mas, S., & Rusiñol, M. (2020). Recent trends on methods for the concentration of viruses from water samples. *Current Opinion in Environmental Science and Health*, 16, 7-13. <https://doi.org/10.1016/j.coesh.2020.01.006>
- Bofill-Mas, S., Rusiñol, M., Fernandez-Cassi, X., Carratalà, A., Hundesa, A., & Girones, R. (2013). Quantification of human and animal viruses to differentiate the origin of the fecal contamination present in environmental samples. *BioMed Research International*, 2013. <https://doi.org/10.1155/2013/192089>
- Borchardt, M. A., Boehm, A. B., Salit, M., Spencer, S. K., Wigginton, K. R., & Noble, R. T. (2021). The Environmental Microbiology Minimum Information (EMMI) Guidelines: QPCR and dPCR Quality and Reporting for Environmental Microbiology. *Environmental Science and Technology*, 55(15), 10210-10223. <https://doi.org/10.1021/acs.est.1c01767>
- Boucherabine, S., Nassar, R., Zaher, S., Mohamed, L., Olsen, M., Alqutami, F., ... Senok, A. (2022). Metagenomic Sequencing and Reverse Transcriptase PCR Reveal That Mobile Phones and Environmental Surfaces Are Reservoirs of Multidrug-Resistant Superbugs and SARS-CoV-2. *Frontiers in Cellular and Infection Microbiology*, 12(March), 1-11. <https://doi.org/10.3389/fcimb.2022.806077>

- Bouseettine, R., Hassou, N., Bessi, H., & Ennaji, M. M. (2020). Waterborne transmission of enteric viruses and their impact on public health. *Emerging and Reemerging Viral Pathogens: Volume 1: Fundamental and Basic Virology Aspects of Human, Animal and Plant Pathogens*, 907-932. <https://doi.org/10.1016/B978-0-12-819400-3.00040-5>
- Breitbart, M., & Rohwer, F. (2005). Here a virus, there a virus, everywhere the same virus? *Trends in Microbiology*, 13(6), 278-284. <https://doi.org/10.1016/j.tim.2005.04.003>
- Briese, T., Kapoor, A., Mishra, N., Jain, K., Kumar, A., Jabado, O. J., & Lipkin, W. I. (2015). Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *mBio*, 6(5). <https://doi.org/10.1128/mBio.01491-15>
- Brooks, J. P., Tanner, B. D., Josephson, K. L., Gerba, C. P., Haas, C. N., & Pepper, I. L. (2005). A national study on the residential impact of biological aerosols from the land application of biosolids. *Journal of Applied Microbiology*, 99(2), 310-322. <https://doi.org/10.1111/j.1365-2672.2005.02604.x>
- Brooks, J. P., McLaughlin, M. R., Gerba, C. P., & Pepper, I. L. (2012). Land Application of Manure and Class B Biosolids: An Occupational and Public Quantitative Microbial Risk Assessment. *Journal of Environmental Quality*, 41(6), 2009-2023. <https://doi.org/10.2134/jeq2011.0430>
- Brum, J. R., Ignacio-espinoza, J. C., Roux, S., Doulcier, G., Acinas, S. G., Alberti, A., & Chaffron, S. (2015). Ocean Viral Communities. *Science*, 348(6237), 1261498-1-11.
- Bustin, S. A., Benes, V., Garson, J. A., Helleman, J., Huggett, J., Kubista, M., ... Wittwer, C. T. (2009). The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, 55(4), 611-622. <https://doi.org/10.1373/clinchem.2008.112797>

REFERÈNCIES

- Callejón, R. M., Rodríguez-Naranjo, M. I., Ubeda, C., Hornedo-Ortega, R., Garcia-Parrilla, M. C., & Troncoso, A. M. (2015). Reported foodborne outbreaks due to fresh produce in the united states and European Union: Trends and causes. *Foodborne Pathogens and Disease*, 12(1), 32-38.
<https://doi.org/10.1089/fpd.2014.1821>
- Cantalupo, P. G., Calgua, B., Zhao, G., Hundesa, A., Wier, A. D., Katz, J. P., ... Pipas, J. M. (2011). Raw sewage harbors diverse viral populations. *mBio*, 2(5). <https://doi.org/10.1128/mBio.00180-11>
- Carducci, A., Donzelli, G., Cioni, L., Federigi, I., Lombardi, R., & Verani, M. (2018). Quantitative microbial risk assessment for workers exposed to bioaerosol in wastewater treatment plants aimed at the choice and setup of safety measures. *International Journal of Environmental Research and Public Health*, 15(1490). <https://doi.org/10.3390/ijerph15071490>
- Carducci, A., Donzelli, G., Cioni, L., & Verani, M. (2016). Quantitative microbial risk assessment in occupational settings applied to the airborne human adenovirus infection. *International Journal of Environmental Research and Public Health*, 13(733). <https://doi.org/10.3390/ijerph13070733>
- Carratalà, A., Rusiñol, M., Hundesa, A., Biarnes, M., Rodriguez-Manzano, J., Vantarakis, A., ... Bofill-Mas, S. (2012). A novel tool for specific detection and quantification of chicken/Turkey parvoviruses to trace poultry fecal contamination in the environment. *Applied and Environmental Microbiology*, 78(20), 7496-7499. <https://doi.org/10.1128/AEM.01283-12>
- CDC, Centers for Disease Control and Prevention (2023). *Norovirus Burden and Trends* [en línia]. Disponible a: <https://www.cdc.gov/norovirus/burden.html> [Últim accés: 09/04/2024].
- Ceballos-Garzon, A., Comtet-Marre, S., & Peyret, P. (2024). Applying targeted gene hybridization capture to viruses with a focus to SARS-CoV-2. *Virus Research*, 340(November 2023), 199293.
<https://doi.org/10.1016/j.virusres.2023.199293>

- Cela-Dablanca, R., Santás-Miguel, V., Fernández-Calviño, D., Arias-Estévez, M., Fernández-Sanjurjo, M. J., Álvarez-Rodríguez, E., & Núñez-Delgado, A. (2021). SARS-CoV-2 and other main pathogenic microorganisms in the environment: Situation in Galicia and Spain. *Environmental Research*, 197(February), 0-3. <https://doi.org/10.1016/j.envres.2021.111049>
- Chappell, J. D., & Dermody, T. S. (2015). Biology of Viruses and Viral Diseases. En *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases* (Eighth Edi, Vol. 2). Elsevier Inc. <https://doi.org/10.1016/B978-1-4557-4801-3.00134-X>
- Corpuz, M. V. A., Buonerba, A., Vigliotta, G., Zarra, T., Ballesteros, F., Campiglia, P., ... Naddeo, V. (2020). Viruses in wastewater: occurrence, abundance and detection methods. *Science of the Total Environment*, 745(140910). <https://doi.org/10.1016/j.scitotenv.2020.140910>
- Corrao, C. R. N., Mazzotta, A., La Torre, G., & De Giusti, M. (2012). Biological risk and occupational Health. *Industrial Health*, 50(4), 326-337. <https://doi.org/10.2486/indhealth.MS1324>
- Corzo, C. A., Culhane, M., Dee, S., Morrison, R. B., & Torremorell, M. (2013). Airborne Detection and Quantification of Swine Influenza A Virus in Air Samples Collected Inside, Outside and Downwind from Swine Barns. *PLoS ONE*, 8(8), 1-7. <https://doi.org/10.1371/journal.pone.0071444>
- Dada, A. C., & Gyawali, P. (2021). Quantitative microbial risk assessment (QMRA) of occupational exposure to SARS-CoV-2 in wastewater treatment plants. *Science of The Total Environment*, 763(142989). <https://doi.org/10.1016/j.scitotenv.2020.142989>
- de Graaf, M., Van Beek, J., Vennema, H., Podkolzin, A. T., Hewitt, J., Bucardo, F., ... Koopmans, M. P. (2015). Emergence of a novel GII.17 norovirus – end of the GII.4 era? *Eurosurveillance*, 20(26). <https://doi.org/10.2807/1560-7917.es2015.20.26.21178>

REFERÈNCIES

- de Jonge, E. F., Peterse, C. M., Koelewijn, J. M., van der Drift, A. M. R., van der Beek, R. F. H. J., Nagelkerke, E., & Lodder, W. J. (2022). The detection of monkeypox virus DNA in wastewater samples in the Netherlands. *Science of the Total Environment*, 852(August), 48-51.
<https://doi.org/10.1016/j.scitotenv.2022.158265>
- Dee, S., Otake, S., Oliveira, S., & Deen, J. (2009). Evidence of long distance airborne transport of porcine reproductive and respiratory syndrome virus and *Mycoplasma hyopneumoniae*. *Veterinary Research*, 40(4).
<https://doi.org/10.1051/vetres/2009022>
- Deere, D., & Ryan, U. (2022). Current assumptions for quantitative microbial risk assessment (QMRA) of Norovirus contamination of drinking water catchments due to recreational activities: an update. *Journal of Water and Health*, 20(10), 1543-1557. <https://doi.org/10.2166/wh.2022.114>
- Desdouits, M., Wacrenier, C., Ollivier, J., Schaeffer, J., & Le Guyader, F. S. (2020). A targeted metagenomics approach to study the diversity of norovirus GII in shellfish implicated in outbreaks. *Viruses*, 12, 978.
<https://doi.org/10.3390/v12090978>
- Di Bonito, P., Della Libera, S., Petricca, S., Iaconelli, M., Sanguinetti, M., Graffeo, R., ... La Rosa, G. (2015). A large spectrum of alpha and beta papillomaviruses are detected in human stool samples. *Journal of General Virology*, 96, 607-613. <https://doi.org/10.1099/vir.0.071787-0>
- Djikeng, A., Kuzmickas, R., Anderson, N. G., & Spiro, D. J. (2009). Metagenomic analysis of RNA viruses in a fresh water lake. *PLoS ONE*, 4(9).
<https://doi.org/10.1371/journal.pone.0007264>
- Donham, K. J., Reynolds, S. J., Whitten, P., Merchant, J. A., Burmeister, L. F., & Popendorf, W. J. (1995). Respiratory Dysfunction in Swine Production Facility Workers: Dose-Response Relationships of Environmental Exposures and Pulmonary Function. *American Journal of Industrial Medicine*, 27, 405-418. <https://doi.org/10.1002/ajim.4700270309>

- Douwes, J., Thorne, P., Pearce, N., & Heederik, D. (2003). Bioaerosol health effects and exposure assessment: Progress and prospects. *Annals of Occupational Hygiene*, 47(3), 187-200.
<https://doi.org/10.1093/annhyg/meg032>
- Douwes, J., Mannelje, A., & Heederik, D. (2001). Work-related symptoms in sewage treatment workers. *Ann Agric Environ Med*, 8, 39-45.
- Duhaime, M. B., Deng, L., Poulos, B. T., & Sullivan, M. B. (2012). Towards quantitative metagenomics of wild viruses and other ultra-low concentration DNA samples: A rigorous assessment and optimization of the linker amplification method. *Environmental Microbiology*, 14(9), 2526-2537.
<https://doi.org/10.1111/j.1462-2920.2012.02791.x>
- EPA, United States Environmental Protection Agency (2023). *What is an emerging viral pathogen claim?* [en línia]. Disponible a:
<https://www.epa.gov/coronavirus/what-emerging-viral-pathogen-claim>
[Últim accés: 09/04/2024].
- Ermonval, M., & Morand, S. (2023). Viral Zoonoses: Interactions and Factors Driving Virus Transmission. *Viruses*, 16(1), 9.
<https://doi.org/10.3390/v16010009>
- Fazil, A. M. (2005). A primer on risk assessment modelling: focus on seafood products. *FAO Fisheries Technical Paper*, 462, 56.
- Fernandez-Cassi, X., Timoneda, N., Gonzales-Gustavson, E., Abril, J. F., Bofill-Mas, S., & Girones, R. (2017). A metagenomic assessment of viral contamination on fresh parsley plants irrigated with fecally tainted river water. *International Journal of Food Microbiology*, 257(December 2016), 80-90. <https://doi.org/10.1016/j.ijfoodmicro.2017.06.001>
- Fernandez-Cassi, X., Timoneda, N., Martínez-Puchol, S., Rusiñol, M., Rodríguez-Manzano, J., Figuerola, N., ... Girones, R. (2018). Metagenomics for the study of viruses in urban sewage as a tool for public health surveillance. *Science of the Total Environment*, 618, 870-880.
<https://doi.org/10.1016/j.scitotenv.2017.08.249>

REFERÈNCIES

- Fuhrimann, S., Winkler, M. S., Stalder, M., Niwagaba, C. B., Babu, M., Kabatereine, N. B., ... Nauta, M. (2016). Disease burden due to gastrointestinal pathogens in a wastewater system in Kampala, Uganda. *Microbial Risk Analysis*, 4, 16-28.
<https://doi.org/10.1016/j.mran.2016.11.003>
- Garcia, B. C. B., Dimasupil, M. A. A. Z., Vital, P. G., Widmer, K. W., & Rivera, W. L. (2015). Fecal contamination in irrigation water and microbial quality of vegetable primary production in urban farms of Metro Manila, Philippines. *J. Environ. Sci. Health.*, 50, 734-743.
<https://doi.org/10.1080/03601234.2015.1048107>
- Gencat, Generalitat de Catalunya - Canal Salut (2024). COVID-19 [en línia]. Disponible a: <https://canalsalut.gencat.cat/ca/salut-a-z/c/covid-19/> [Últim accés: 09/04/2024].
- Gencat, Generalitat de Catalunya - Scientia. Dipòsit d'Informació Digital del Departament de Salut (2023). *Situació epidemiològica dels casos d'Mpox (verola del mico) a Catalunya* [en línia]. Disponible a: <https://scientiasalut.gencat.cat/handle/11351/10691> [Últim accés: 09/04/2024].
- Geng-Hao, B., Sheng-Chieh, L., Yi-Hsiang, H., & Shih-Yen, C. (2022). The Human Virome: Viral Metagenomics, Relations with Human Diseases, and Therapeutic Applications. *Viruses*, 14(278), 1-29.
- Gholipour, S., Mohammadi, F., Nikaeen, M., Shamsizadeh, Z., Khazeni, A., Sahbaei, Z., ... Mirhendi, H. (2021). COVID-19 infection risk from exposure to aerosols of wastewater treatment plants. *Chemosphere*, 273, 129701.
<https://doi.org/10.1016/j.chemosphere.2021.129701>
- Gibson, K. E. (2014). Viral pathogens in water: Occurrence, public health impact, and available control strategies. *Current Opinion in Virology*, 4, 50-57.
<https://doi.org/10.1016/j.coviro.2013.12.005>

- Girardi, V., Mena, K. D., Albino, S. M., Demoliner, M., Gularte, J. S., de Souza, F. ... Spilki, F. R. (2019). Microbial risk assessment in recreational freshwaters from southern Brazil. *Science of the Total Environment*, 651, 298-308. <https://doi.org/10.1016/j.scitotenv.2018.09.177>
- Gobeil, A., Maherani, B., & Lacroix, M. (2020). Norovirus elimination on the surface of fresh foods. *Critical Reviews in Food Science and Nutrition*. <https://doi.org/10.1080/10408398.2020.1848784>
- Gonzales-Gustavson, E., Rusiñol, M., Medema, G., Calvo, M., & Girones, R. (2019). Quantitative risk assessment of norovirus and adenovirus for the use of reclaimed water to irrigate lettuce in Catalonia. *Water Research*, 153, 91-99. <https://doi.org/10.1016/j.watres.2018.12.070>
- Gordon-Lipkin, E. M., Marcum, C. S., Kruk, S., Thompson, E., Kelly, S. E. M., Kalish, H., ... McGuire, P. J. (2022). Comprehensive profiling of the human viral exposome in households containing an at-risk child with mitochondrial disease during the 2020–2021 COVID-19 pandemic. *Clinical and Translational Medicine*, 12(11). <https://doi.org/10.1002/ctm2.1100>
- Grafanaki, K., Bania, A., Kaliatsi, E. G., Vryzaki, E., Vasilopoulos, Y., & Georgiou, S. (2023). The Imprint of Exposome on the Development of Atopic Dermatitis across the Lifespan: A Narrative Review. *Journal of Clinical Medicine*, 12(6). <https://doi.org/10.3390/jcm12062180>
- Gray, G. C., McCarthy, T., Capuano, A. W., Setterquist, S. F., Olsen, C. W., Alavanja, M. C., & Lynch, C. F. (2007). Swine workers and swine influenza virus infections. *Emerging Infectious Diseases*, 13(12), 1871-1878. <https://doi.org/10.3201/eid1312.061323>
- Gularte, J. S., Girardi, V., Demoliner, M., de Souza, F. G., Filippi, M., Eisen, A. K. ... Spilki, F. R. (2019). Human mastadenovirus in water, sediment, sea surface microlayer, and bivalve mollusk from southern Brazilian beaches. *Marine Pollution Bulletin*, 142(December 2018), 335-349. <https://doi.org/10.1016/j.marpolbul.2018.12.046>

REFERÈNCIES

- Haas, C. N. (1983a). Effect of effluent disinfection on risks of viral disease transmission via recreational water exposure. *Journal of the Water Pollution Control Federation*, 55(8), 1111-1116.
- Haas, C. N. (1983b). Estimation of risk due to low doses of microorganisms: A comparison of alternative methodologies. *American Journal of Epidemiology*, 118(4), 573-582.
<https://doi.org/10.1093/oxfordjournals.aje.a113662>
- Haas, C. N. (2020). Quantitative Microbial Risk Assessment and Molecular Biology: Paths to Integration. *Environmental Science and Technology*, 54(14), 8539-8546. <https://doi.org/10.1021/acs.est.0c00664>
- Haas, C. N., Rose, J. B., & Gerba, C. P. (2014). Risk Assessment Paradigms. En *Quantitative Microbial Risk Assessment* (p. 63-89). Wiley.
<https://doi.org/10.1002/9781118910030.ch3>
- Hall, R. J., Leblanc-Maridor, M., Wang, J., Ren, X., Moore, N. E., Brooks, C. R., ... McLean, D. J. (2013). Metagenomic Detection of Viruses in Aerosol Samples from Workers in Animal Slaughterhouses. *PLoS ONE*, 8(8).
<https://doi.org/10.1371/journal.pone.0072226>
- Hamza, H., & Hamza, I. A. (2018). Oncogenic papillomavirus and polyomavirus in urban sewage in Egypt. *Science of the Total Environment*, 610-611, 1413-1420. <https://doi.org/10.1016/j.scitotenv.2017.08.218>
- Hamza, I. A., & Bibby, K. (2019). Critical issues in application of molecular methods to environmental virology. *Journal of Virological Methods*, 266(January), 11-24. <https://doi.org/10.1016/j.jviromet.2019.01.008>
- Han, J. J., Song, H. A., Pierson, S. L., Shen-Gunther, J., & Xia, Q. (2023). Emerging Infectious Diseases Are Virulent Viruses—Are We Prepared? An Overview. *Microorganisms*, 11(11), 1-17.
<https://doi.org/10.3390/microorganisms11112618>

- Han, Y., Li, L., & Liu, J. (2013). Characterization of the airborne bacteria community at different distances from the rotating brushes in a wastewater treatment plant by 16S rRNA gene clone libraries. *Journal of Environmental Sciences*, 25(1), 5-15. [https://doi.org/10.1016/S1001-0742\(12\)60018-7](https://doi.org/10.1016/S1001-0742(12)60018-7)
- Han, Y., Yang, T., Chen, T., Li, L., & Liu, J. (2019). Characteristics of submicron aerosols produced during aeration in wastewater treatment. *Science of the Total Environment*, 696, 134019 <https://doi.org/10.1016/j.scitotenv.2019.134019>
- Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J., & Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. *Chemistry and Biology*, 5(10). [https://doi.org/10.1016/S1074-5521\(98\)90108-9](https://doi.org/10.1016/S1074-5521(98)90108-9)
- Haramoto, E., Kitajima, M., Hata, A., Torrey, J. R., Masago, Y., Sano, D., & Katayama, H. (2018). A review on recent progress in the detection methods and prevalence of human enteric viruses in water. *Water Research*, 135, 168-186. <https://doi.org/10.1016/j.watres.2018.02.004>
- Harvey, W. T., Carabelli, A. M., Jackson, B., Gupta, R. K., Thomson, E. C., Harrison, E. M., ... Robertson, D. L. (2021). SARS-CoV-2 variants, spike mutations and immune escape. *Nature Reviews Microbiology*, 19(7), 409-424. <https://doi.org/10.1038/s41579-021-00573-0>
- He, W. T., Hou, X., Zhao, J., Sun, J., He, H., Si, W., ... Su, S. (2022). Virome characterization of game animals in China reveals a spectrum of emerging pathogens. *Cell*, 185(7), 1117-1129.e8. <https://doi.org/10.1016/j.cell.2022.02.014>
- Head, S. R., Kiyomi Komori, H., LaMere, S. A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D. R., & Ordoukhanian, P. (2014). Library construction for next-generation sequencing: Overviews and challenges. *BioTechniques*, 56(2), 61-77. <https://doi.org/10.2144/000114133>

- Hellmér, M., Paxéus, N., Magnius, L., Enache, L., Arnholm, B., Johansson, A., ... Norder, H. (2014). Detection of pathogenic viruses in sewage provided early warnings of hepatitis A virus and norovirus outbreaks. *Applied and Environmental Microbiology*, 80(21), 6771-6781.
<https://doi.org/10.1128/AEM.01981-14>
- Hijnen, W. A. M., Beerendonk, E. F., & Medema, G. J. (2006). Inactivation credit of UV radiation for viruses, bacteria and protozoan (oo)cysts in water: A review. *Water Research*, 40(1), 3-22.
<https://doi.org/10.1016/j.watres.2005.10.030>
- Hjelmso, M. H., Mollerup, S., Jensen, R. H., Pietroni, C., Lukjancenko, O., Schultz, ... Hansen, A. J. (2019). Metagenomic analysis of viruses in toilet waste from long distance flights—A new procedure for global infectious disease surveillance. *PLoS ONE*, 14(1).
<https://doi.org/10.1371/journal.pone.0210368>
- Hood, G., Roche, X., Brioude, A., von Dobschuetz, S., Fasina, F. O., Kalpravidh, W., ... Sims, L. (2021). A literature review of the use of environmental sampling in the surveillance of avian influenza viruses. *Transboundary and Emerging Diseases*, 68(1), 110-126.
<https://doi.org/10.1111/tbed.13633>
- Hsu, T., Joice, R., Vallarino, J., Abu-Ali, G., Hartmann, E. M., Shafquat, A., ... Huttenhower, C. (2016). Urban Transit System Microbial Communities Differ by Surface Type and Interaction with Humans and the Environment. *mSystems*, 1(3). <https://doi.org/10.1128/msystems.00018-16>
- Hu, L., Deng, W. J., Ying, G. G., & Hong, H. (2021). Environmental perspective of COVID-19: Atmospheric and wastewater environment in relation to pandemic. *Ecotoxicology and Environmental Safety*, 219(January), 112297.
<https://doi.org/10.1016/j.ecoenv.2021.112297>
- Huang, W., Bai, L., & Tang, H. (2023). Epstein-Barr virus infection: the micro and macro worlds. *Virology Journal*, 20(220). <https://doi.org/10.1186/s12985-023-02187-9>

- Hurwitz, B. L., Deng, L., Poulos, B. T., & Sullivan, M. B. (2013). Evaluation of methods to concentrate and purify ocean virus communities through comparative, replicated metagenomics. *Environmental Microbiology*, 15(5), 1428-1440. <https://doi.org/10.1111/j.1462-2920.2012.02836.x>
- Iaconelli, M., Petricca, S., Libera, S. Della, Di Bonito, P., & La Rosa, G. (2015). First Detection of Human Papillomaviruses and Human Polyomaviruses in River Waters in Italy. *Food and Environmental Virology*, 7(4), 309-315. <https://doi.org/10.1007/s12560-015-9203-7>
- Iaconelli, M., Ferraro, G. B., Mancini, P., Suffredini, E., Veneri, C., Ciccaglione, A. R., ... Rosa, G. La. (2020). Nine-year nationwide environmental surveillance of hepatitis E virus in urban wastewaters in Italy (2011–2019). *International Journal of Environmental Research and Public Health*, 17(6), 1-18. <https://doi.org/10.3390/ijerph17062059>
- Idescat, Institut d'Estadística de Catalunya (2022). *Qualitat de l'aigua. Controls efectuats en aigües continentals i costaneres de Catalunya* [en línia]. Disponible a: <https://www.idescat.cat/indicadors/?id=aec&n=15210> [Últim accés: 09/04/2024]
- Imamura, S., Kanezashi, H., Goshima, T., Haruna, M., Okada, T., Inagaki, N., ... Akimoto, K. (2017). Next-Generation Sequencing Analysis of the Diversity of Human Noroviruses in Japanese Oysters. *Foodborne Pathogens and Disease*, 14(8), 465-471. <https://doi.org/10.1089/fpd.2017.2289>
- Islam, A., Ferdous, J., Islam, S., Sayeed, M. A., Rahman, M. K., Saha, O., ... Shirin, T. (2022). Transmission dynamics and susceptibility patterns of SARS-CoV-2 in domestic, farmed and wild animals: Sustainable One Health surveillance for conservation and public health to prevent future epidemics and pandemics. *Transboundary and Emerging Diseases*, 69(5), 2523-2543. <https://doi.org/10.1111/tbed.14356>

REFERÈNCIES

- ISO, International Organization for Standardization (2017). *ISO 15216-1:2017. Microbiology of the food chain — Horizontal method for determination of hepatitis A virus and norovirus in food using real-time RT-PCR — Part 1: Method for quantification* [en línia].
Disponible a: <https://www.iso.org/standard/65681.html> [Últim accés: 09/04/2024].
- Kalantar, K. L., Carvalho, T., De Bourcy, C. F. A., Dimitrov, B., Dingle, G., Egger, R., ... DeRisi, J. L. (2020). IDseq-An open source cloud-based pipeline and analysis service for metagenomic pathogen detection and monitoring. *GigaScience*, 9, 1-14. <https://doi.org/10.1093/gigascience/giaa111>
- Kallies, R., Hölzer, M., Toscan, R. B., da Rocha, U. N., Anders, J., Marz, M., & Chatzinotas, A. (2019). Evaluation of sequencing library preparation protocols for viral metagenomic analysis from pristine aquifer groundwaters. *Viruses*, 11(6). <https://doi.org/10.3390/v11060484>
- Karlsson, O. E., Belák, S., & Granberg, F. (2013). *The Effect of Preprocessing by Sequence-Independent, Single-Primer Amplification (SISPA) on Metagenomic Detection of Viruses*. 11, S227-S234
<https://doi.org/10.1089/bsp.2013.0008>
- Kataki, S., Patowary, R., Chatterjee, S., Vairale, M. G., Sharma, S., Dwivedi, S. K., & Kamboj, D. V. (2022). Bioaerosolization and pathogen transmission in wastewater treatment plants: Microbial composition, emission rate, factors affecting and control measures. *Chemosphere*, 287(P3), 132180.
<https://doi.org/10.1016/j.chemosphere.2021.132180>
- Kitakawa, K., Kitamura, K., & Yoshida, H. (2023). Monitoring Enteroviruses and SARS-CoV-2 in Wastewater Using the Polio Environmental Surveillance System in Japan. *Applied and Environmental Microbiology*, 89(4).
<https://doi.org/10.1128/aem.01853-22>
- Kocjan, B. J., Seme, K., Cimerman, M., Kovanda, A., Potočnik, M., & Poljak, M. (2009). Genomic diversity of human papillomavirus (HPV) Genotype 38. *Journal of Medical Virology*, 81(2), 288-295
<https://doi.org/10.1002/jmv.21392>

- Koo, H. L., Ajami, N., Atmar, R. L., & DuPont, H. L. (2010). Noroviruses: The leading cause of gastroenteritis worldwide. *Discov. Med.*, 10(50), 61-70.
- Koonin, E. V., Dolja, V. V., & Krupovic, M. (2021). The healthy human virome: from virus–host symbiosis to disease. *Current Opinion in Virology*, 47, 86-94. <https://doi.org/10.1016/j.coviro.2021.02.002>
- Koopmans, M., & Duizer, E. (2004). Foodborne viruses: An emerging problem. *International Journal of Food Microbiology*, 90(1), 23-41. [https://doi.org/10.1016/S0168-1605\(03\)00169-7](https://doi.org/10.1016/S0168-1605(03)00169-7)
- Kuchinski, K. S., Loos, K. D., Suchan, D. M., Russell, J. N., Sies, A. N., Kumakamba, C., ... Cameron, A. D. (2022). Targeted genomic sequencing with probe capture for discovery and surveillance of coronaviruses in bats. *eLife*, 11, 1-25. <https://doi.org/10.7554/ELIFE.79777>
- Kumar, A., Murthy, S., & Kapoor, A. (2017). Evolution of selective-sequencing approaches for virus discovery and virome analysis. *Virus Research*, 239(September 2016), 172-179. <https://doi.org/10.1016/j.virusres.2017.06.005>
- La Rosa, G., Della Libera, S., Petricca, S., Iaconelli, M., Briancesco, R., Paradiso, R., ... Bonadonna, L. (2015). First detection of papillomaviruses and polyomaviruses in swimming pool waters: Unrecognized recreational water-related pathogens? *Journal of Applied Microbiology*, 119(6), 1683-1691. <https://doi.org/10.1111/jam.12925>
- La Rosa, G., Fratini, M., Accardi, L., D'Oro, G., Della Libera, S., Muscillo, M., & Di Bonito, P. (2013). Mucosal and Cutaneous Human Papillomaviruses Detected in Raw Sewages. *PLoS ONE*, 8(1), e52391. <https://doi.org/10.1371/journal.pone.0052391>
- La Rosa, G., Fratini, M., Della Libera, S., Iaconelli, M., & Muscillo, M. (2012). Emerging and potentially emerging viruses. *Annali dell'Istituto Superiore di Sanità*, 48, 397-406.

REFERÈNCIES

- Labadie, T., Batéjat, C., Leclercq, I., & Manuguerra, J. C. (2020). Historical Discoveries on Viruses in the Environment and Their Impact on Public Health. *Intervirology*, 63(1-6), 17-32. <https://doi.org/10.1159/000511575>
- Landrigan, P. J., Stegeman, J. J., Fleming, L. E., Allemand, D., Anderson, D. M., Backer, L. C., ... Rampal, P. (2020). Human health and ocean pollution. *Annals of Global Health*, 86(1), 1-64. <https://doi.org/10.5334/aogh.2831>
- Lanzarini, N. M., Federigi, I., Marinho Mata, R., Neves Borges, M. D., Mendes Saggioro, E., ... Pereira Miagostovich, M. (2022). Human adenovirus in municipal solid waste leachate and quantitative risk assessment of gastrointestinal illness to waste collectors. *Waste Management*, 138(August 2021), 308-317. <https://doi.org/10.1016/j.wasman.2021.12.005>
- Le Guyader, F. S., Mittelholzer, C., Haugarreau, L., Hedlund, K. O., Alsterlund, R., Pommepuy, M., & Svensson, L. (2004). Detection of noroviruses in raspberries associated with a gastroenteritis outbreak. *International Journal of Food Microbiology*, 97(2), 179-186. <https://doi.org/10.1016/j.ijfoodmicro.2004.04.018>
- Lecuit, M., & Eloit, M. (2013). The human virome: New tools and concepts. *Trends in Microbiology*, 21(10), 510-515. <https://doi.org/10.1016/j.tim.2013.07.001>
- Lees, D. (2000). Viruses and bivalve shellfish. *International Journal of Food Microbiology*, 59(1-2), 81-116. [https://doi.org/10.1016/S0168-1605\(00\)00248-8](https://doi.org/10.1016/S0168-1605(00)00248-8)
- Li, J., Pan, Y. Q., Xu, Z., Wang, Q., Hang, D., Shen, N., ... Ke, Y. (2013). Improved detection of human papillomavirus harbored in healthy skin with FAP6085/64 primers. *Journal of Virological Methods*, 193(2), 633-638. <https://doi.org/10.1016/j.jviromet.2013.06.026>
- Liu, T., Li, G., Liu, Z., Xi, L., Ma, W., & Gao, X. (2023). Characteristics of aerosols from swine farms: A review of the past two-decade progress. *Environment International*, 178(June), 108074. <https://doi.org/10.1016/j.envint.2023.108074>

- Liu, W., MacDonald, M., & You, J. (2016). Merkel cell polyomavirus infection and Merkel cell carcinoma. *Current Opinion in Virology*, 20, 20-27.
<https://doi.org/10.1016/j.coviro.2016.07.011>
- Louten, J. (2016). Virus Transmission and Epidemiology. En *Essential Human Virology* (Vol. 14, Número 2, p. 71-92). Elsevier.
<https://doi.org/10.1016/B978-0-12-800947-5.00005-3>
- Macdiarmid, S. C., & Pharo, H. J. (2003). Risk analysis: assessment , management and How should risk analysis be conducted? Hazard identification. *Revue scientifique et technique (International Office of Epizootics)*, 22(2), 397-408.
- Mackul'ak, T., Gál, M., Špalková, V., Fehér, M., Briestenská, K., Mikušová, M., ... Škulcová, A. B. (2021). Wastewater-based epidemiology as an early warning system for the spreading of sars-cov-2 and its mutations in the population. *International Journal of Environmental Research and Public Health*, 18(11). <https://doi.org/10.3390/ijerph18115629>
- Mainelis, G. (2020). Bioaerosol sampling: Classical approaches, advances, and perspectives. *Aerosol Science and Technology*, 54(5), 496-519.
<https://doi.org/10.1080/02786826.2019.1671950>
- Mandal, J., & Brandl, H. (2011). Bioaerosols in Indoor Environment - A Review with Special Reference to Residential and Occupational Locations. *The Open Environmental & Biological Monitoring Journal*, 4(1), 83-96.
<https://doi.org/10.2174/1875040001104010083>
- Martínez-Puchol, S., Rusiñol, M., Fernández-Cassi, X., Timoneda, N., Itarte, M., Andrés, C., ... Bofill-Mas, S. (2020a). Characterisation of the sewage virome: comparison of NGS tools and occurrence of significant pathogens. *Science of the Total Environment*, 713.
<https://doi.org/10.1016/j.scitotenv.2020.136604>
- Martínez-Puchol, S. (2020b). *Seqüenciació massiva aplicada a l'epidemiologia d'aigües residuals i a la caracterització de viromes*. Tesis doctoral: Microbiologia Ambiental i Biotecnologia. Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona, Barcelona.

REFERÈNCIES

- Masclaux, F. G., Hotz, P., Gashi, D., Savova-Bianchi, D., & Oppliger, A. (2014). Assessment of airborne virus contamination in wastewater treatment plants. *Environmental Research*, 133, 260-265.
<https://doi.org/10.1016/j.envres.2014.06.002>
- Medema, G., Wuilings, B., Roeleveld, P., & Van Der Kooij, D. (2004). Risk assessment of Legionella and enteric pathogens in sewage treatment works. *Water Science and Technology: Water Supply*, 4(2), 125-132.
<https://doi.org/10.2166/ws.2004.0037>
- Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., & Brouwer, A. (2020). Presence of SARS-Coronavirus-2 RNA in Sewage and Correlation with Reported COVID-19 Prevalence in the Early Stage of the Epidemic in the Netherlands. *Environmental Science and Technology Letters*, 7(7), 511-516.
<https://doi.org/10.1021/acs.estlett.0c00357>
- Mejías-Molina, C., Pico-Tomás, A., Beltran-Rubinat, A., Martínez-Puchol, S., Corominas, L., Rusiñol, M., & Bofill-Mas, S. (2023). Effectiveness of passive sampling for the detection and genetic characterization of human viruses in wastewater. *Environmental Science: Water Research and Technology*.
<https://doi.org/10.1039/d2ew00867j>
- Mirskaya, E., & Agranovski, I. E. (2018). Sources and mechanisms of bioaerosol generation in occupational environments. *Critical Reviews in Microbiology*, 44(6), 739-758. <https://doi.org/10.1080/1040841X.2018.1508125>
- Myers, K. P., Olsen, C. W., Setterquist, S. F., Capuano, A. W., Donham, K. J., Thacker, E. L., ... Gray, G. C. (2006). Are swine workers in the United States at increased risk of infection with zoonotic influenza virus? *Clinical Infectious Diseases*, 42(1), 14-20. <https://doi.org/10.1086/498977>
- NCBI, National Center for Biotechnology Information (2013). *GenBank* [en línia]. Disponible a:
<https://www.ncbi.nlm.nih.gov/genbank/> [Últim accés: 09/04/2024]

- Neira, V., Rabinowitz, P., Rendahl, A., Paccha, B., Gibbs, S. G., & Torremorell, M. (2016). Characterization of viral load, viability and persistence of influenza A virus in air and on surfaces of swine production facilities. *PLoS ONE*, 11(1), 1-11. <https://doi.org/10.1371/journal.pone.0146616>
- Newell, D. G., Koopmans, M., Verhoef, L., Duizer, E., Aidara-Kane, A., Sprong, H., ... Kruse, H. (2010). Food-borne diseases — The challenges of 20 years ago still persist while new ones continue to emerge. *International Journal of Food Microbiology*, 139(SUPPL. 1), S3-S15. <https://doi.org/10.1016/j.ijfoodmicro.2010.01.021>
- Núñez-Delgado, A. (2020). What do we know about the SARS-CoV-2 coronavirus in the environment? *Science of the Total Environment*, 727, 138647. <https://doi.org/10.1016/j.scitotenv.2020.138647>
- Nwachuku, N., & Gerba, C. P. (2004). Microbial risk assessment: Don't forget the children. *Current Opinion in Microbiology*, 7(3), 206-209. <https://doi.org/10.1016/j.mib.2004.04.011>
- Ogorzaly, L., Walczak, C., Galloux, M., Etienne, S., Gassilloud, B., & Cauchie, H. M. (2015). Human Adenovirus Diversity in Water Samples Using a Next-Generation Amplicon Sequencing Approach. *Food and Environmental Virology*, 7, 112-121. <https://doi.org/10.1007/s12560-015-9194-4>
- Oshiki, M., Miura, T., Kazama, S., Segawa, T., Ishii, S., Hatamoto, M., ... Sano, D. (2018). Microfluidic PCR amplification and MiSeq amplicon sequencing techniques for high-throughput detection and genotyping of human pathogenic RNA viruses in human feces, sewage, and oysters. *Frontiers in Microbiology*, 9, 830. <https://doi.org/10.3389/fmicb.2018.00830>
- Owens, C. E. L., Angles, M. L., Cox, P. T., Byleveld, P. M., Osborne, N. J., & Rahman, M. B. (2020). Implementation of quantitative microbial risk assessment (QMRA) for public drinking water supplies: Systematic review. *Water Research*, 174, 115614. <https://doi.org/10.1016/j.watres.2020.115614>

REFERÈNCIES

- Paccha, B., Jones, R. M., Gibbs, S., Kane, M. J., Torremorell, M., Neira-Ramirez, V., & Rabinowitz, P. M. (2016). Modeling risk of occupational zoonotic influenza infection in swine workers. *Journal of Occupational and Environmental Hygiene*, 13(8), 577-587.
<https://doi.org/10.1080/15459624.2016.1159688>
- Pachepsky, Y., Shelton, D. R., McLain, J. E. T., Patel, J., & Mandrell, R. E. (2011). Irrigation Waters as a Source of Pathogenic Microorganisms in Produce. A Review. *Advances in Agronomy*, 113, 73-138. <https://doi.org/10.1016/B978-0-12-386473-4.00007-5>
- Pepper, I. L., & Gerba, C. P. (2015). Environmental Sample Collection and Processing. En *Environmental Microbiology: Third Edition*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-394626-3.00008-9>
- Petterson, S. R. (2016). Application of a QMRA Framework to Inform Selection of Drinking Water Interventions in the Developing Context. *Risk Analysis*, 36(2), 203-214. <https://doi.org/10.1111/risa.12452>
- Pina, S., Puig, M., Lucena, F., Jofre, J., & Girones, R. (1998). Viral pollution in the environment and in shellfish: human adenovirus detection by PCR as an index of human viruses. *Applied and environmental microbiology*, 64(9), 3376-3382.
- Pogka, V., Papadopoulou, G., Valiakou, V., Sgouras, D. N., Mentis, A. F., & Karamitros, T. (2022). Targeted Virome Sequencing Enhances Unbiased Detection and Genome Assembly of Known and Emerging Viruses—The Example of SARS-CoV-2. *Viruses*, 14(6).
<https://doi.org/10.3390/v14061272>
- Prevost, B., Lucas, F. S., Ambert-Balay, K., Pothier, P., Moulin, L., & Wurtzer, S. (2015). Deciphering the diversities of astroviruses and noroviruses in wastewater treatment plant effluents by a high-throughput sequencing method. *Applied and Environmental Microbiology*, 81(20), 7215-7222. <https://doi.org/10.1128/AEM.02076-15>

- Radford, A. D., Chapman, D., Dixon, L., & Chantrey, J. (2012). Application of Next-Generation Sequencing Technologies in Virology. *Journal of General Virology*, May, 401-412. <https://doi.org/10.1016/B978-0-12-801498-1.00038-3>
- Riesenfeld, C. S., Schloss, P. D., & Handelsman, J. (2004). Metagenomics: Genomic analysis of microbial communities. *Annual Review of Genetics*, 38, 525-552. <https://doi.org/10.1146/annurev.genet.38.072902.091216>
- Robotto, A., Lembo, D., Quaglino, P., Brizio, E., Polato, D., Civra, A., ... Di Perri, G. (2022). Wastewater-based SARS-CoV-2 environmental monitoring for Piedmont, Italy. *Environmental Research*, 203(January), 111901. <https://doi.org/10.1016/j.envres.2021.111901>
- Rodríguez-Lázaro, D., Cook, N., Ruggeri, F. M., Sellwood, J., Nasser, A., Nascimento, ... van der Poel, W. H. M. (2012). Virus hazards from food, water and other contaminated environments. *FEMS Microbiology Reviews*, 36(4), 786-814. <https://doi.org/10.1111/j.1574-6976.2011.00306.x>
- Rodríguez, R. A., Pepper, I. L., & Gerba, C. P. (2009). Application of PCR-based methods to assess the infectivity of enteric viruses in environmental samples. En *Applied and Environmental Microbiology* (Vol. 75, Número 2, p. 297-307). <https://doi.org/10.1128/AEM.01150-08>
- Roossinck, M. J. (2015). Move Over, Bacteria! Viruses Make Their Mark as Mutualistic Microbial Symbionts. *Journal of Virology*, 89(13), 6532-6535. <https://doi.org/10.1128/jvi.02974-14>
- Rusiñol, M., Zammit, I., Itarte, M., Forés, E., Martínez-Puchol, S., Girones, R., ... Bofill-Mas, S. (2021). Monitoring waves of the COVID-19 pandemic: Inferences from WWTPs of different sizes. *Science of the Total Environment*, 787. <https://doi.org/10.1016/j.scitotenv.2021.147463>
- Rusiñol, M. (2023). Waterborne viruses in urban groundwater environments. *PLOS Water*, 2(8), e0000168. <https://doi.org/10.1371/journal.pwat.0000168>

REFERÈNCIES

- Rusiñol, M., Fernandez-Cassi, X., Hundesa, A., Vieira, C., Kern, A., Eriksson, I., ... Girones, R. (2014). Application of human and animal viral microbial source tracking tools in fresh and marine waters from five different geographical areas. *Water Research*, 59, 119-129.
<https://doi.org/10.1016/j.watres.2014.04.013>
- Rusiñol, M., Fernandez-Cassi, X., Timoneda, N., Carratalà, A., Abril, J. F., Silvera, C., ... Girones, R. (2015). Evidence of viral dissemination and seasonality in a Mediterranean river catchment: Implications for water pollution management. *Journal of Environmental Management*, 159, 58-67.
<https://doi.org/10.1016/j.jenvman.2015.05.019>
- Rusiñol, M., & Girones, R. (2017). Summary of Excreted and Waterborne Viruses. *Global Water Pathogen Project*.
<https://doi.org/10.14321/waterpathogens.19>
- Rusiñol, M., Martínez-Puchol, S., Timoneda, N., Fernández-Cassi, X., Pérez-Cataluña, A., Fernández-Bravo, A., ... Girones, R. (2020). Metagenomic analysis of viruses, bacteria and protozoa in irrigation water. *International Journal of Hygiene and Environmental Health*, 224(December).
<https://doi.org/10.1016/j.ijheh.2019.113440>
- Sales-Ortells, H., Agostini, G., & Medema, G. (2015a). Quantification of waterborne pathogens and associated health risks in urban water. *Environmental Science and Technology*, 49(11), 6943-6952.
<https://doi.org/10.1021/acs.est.5b00625>
- Sales-Ortells, H., Fernandez-Cassi, X., Timoneda, N., Dürig, W., Girones, R., & Medema, G. (2015b). Health risks derived from consumption of lettuces irrigated with tertiary effluent containing norovirus. *Food Research International*, 68, 70-77. <https://doi.org/10.1016/j.foodres.2014.08.018>
- Sales-Ortells, H., & Medema, G. (2014). Screening-level microbial risk assessment of urban water locations: A tool for prioritization. *Environmental Science and Technology*, 48(16), 9780-9789.
<https://doi.org/10.1021/es5020407>

- Samadi, S., Wouters, I. M., & Heederik, D. J. J. (2013). A review of bio-aerosol exposures and associated health effects in veterinary practice. *Annals of Agricultural and Environmental Medicine*, 20(2), 206-221.
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, 74(12), 5463-5467. <https://doi.org/10.1073/pnas.74.12.5463>
- Sanguenza-Acosta, M., & Sandoval-Romero, E. (2018). Epstein-Barr virus and skin. *Anais Brasileiros de Dermatologia*, 93(6), 786-799. <https://doi.org/10.1590/abd1806-4841.20187021>
- Santarpia, J. L., Klug, E., Ravnholdt, A., & Kinahan, S. M. (2023). Environmental sampling for disease surveillance: Recent advances and recommendations for best practice. *Journal of the Air and Waste Management Association*, 73(6), 434-461. <https://doi.org/10.1080/10962247.2023.2197825>
- Sauer, E. P., VandeWalle, J. L., Bootsma, M. J., & McLellan, S. L. (2011). Detection of the human specific Bacteroides genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. *Water Research*, 45(14), 4081-4091. <https://doi.org/10.1016/j.watres.2011.04.049>
- Schang, C., Crosbie, N. D., Nolan, M., Poon, R., Wang, M., Jex, A., ... McCarthy, D. T. (2021). Passive Sampling of SARS-CoV-2 for Wastewater Surveillance. *Environmental Science and Technology*, 55(15), 10432-10441. <https://doi.org/10.1021/acs.est.1c01530>
- Schijven, J., Teunis, P., Suylen, T., Ketelaars, H., Hornstra, L., & Rutjes, S. (2019). QMRA of adenovirus in drinking water at a drinking water treatment plant using UV and chlorine dioxide disinfection. *Water Research*, 158, 34-45. <https://doi.org/10.1016/j.watres.2019.03.090>
- Schoen, M. E., Jahne, M. A., & Garland, J. L. (2018). Human Health Impact of Cross-Connections in Non-Potable Reuse Systems. *Water*, 10(10). <https://doi.org/10.3390/w10101352>

REFERÈNCIES

- Segelhurst, E., Bard, J. E., Pillsbury, A. N., Alam, M. M., Lamb, N. A., Zhu, C., ... Ye, Y. (2023). Robust Performance of SARS-CoV-2 Whole-Genome Sequencing from Wastewater with a Nonselective Virus Concentration Method. *ACS ES and T Water*, 3(4), 954-962.
<https://doi.org/10.1021/acsestwater.2c00456>
- Shah, S., Gwee, S. X. W., Ng, J. Q. X., Lau, N., Koh, J., & Pang, J. (2022). Wastewater surveillance to infer COVID-19 transmission: A systematic review. *Science of the Total Environment*, 804, 150060.
<https://doi.org/10.1016/j.scitotenv.2021.150060>
- Shaheen, M. N. F. (2022). The concept of one health applied to the problem of zoonotic diseases. *Reviews in Medical Virology*, 32(4), 1-14.
<https://doi.org/10.1002/rmv.2326>
- Shaheen, M. N. F., Elmahdy, E. M., & Chawla-Sarkar, M. (2019). Quantitative PCR-based identification of enteric viruses contaminating fresh produce and surface water used for irrigation in Egypt. *Environ. Sci. Pollut. Res.*, 26, 21619-21628. <https://doi.org/10.1007/s11356-019-05435-0>
- Sharma, A., & Gilbert, J. A. (2018). Microbial exposure and human health. *Current Opinion in Microbiology*, 44, 79-87.
<https://doi.org/10.1016/j.mib.2018.08.003>
- Sharun, K., Dhama, K., Pawde, A. M., Gortázar, C., Tiwari, R., Bonilla-Aldana, D. ... Attia, Y. A. (2021). SARS-CoV-2 in animals: potential for unknown reservoir hosts and public health implications. *Veterinary Quarterly*, 41(1), 181-201. <https://doi.org/10.1080/01652176.2021.1921311>
- Sheek-Hussein, M., Alsuwaidi, A. R., Davies, E. A., & Abu-Zidan, F. M. (2023). Monkeypox: A current emergency global health threat. *Turkish Journal of Emergency Medicine*, 23(1), 5-16.
<https://doi.org/10.4103/2452-2473.366487>

- Sidhu, J. P. S., Ahmed, W., Gernjak, W., Aryal, R., McCarthy, D., Palmer, A., ... Toze, S. (2013). Sewage pollution in urban stormwater runoff as evident from the widespread presence of multiple microbial and chemical source tracking markers. *Science of the Total Environment*, 463-464, 488-496. <https://doi.org/10.1016/j.scitotenv.2013.06.020>
- Sidhu, J. P. S., Hodggers, L., Ahmed, W., Chong, M. N., & Toze, S. (2012). Prevalence of human pathogens and indicators in stormwater runoff in Brisbane, Australia. *Water Research*, 46(20), 6652-6660. <https://doi.org/10.1016/j.watres.2012.03.012>
- Simmonds, P., Adams, M. J., Benk, M., Breitbart, M., Brister, J. R., Carstens, E. B.,... Zerbini, F. M. (2017). Consensus statement: Virus taxonomy in the age of metagenomics. *Nature Reviews Microbiology*, 15(3), 161-168. <https://doi.org/10.1038/nrmicro.2016.177>
- Singer, A. C., Thompson, J. R., Filho, C. R. M., Street, R., Li, X., Castiglioni, S., & Thomas, K. V. (2023). A world of wastewater-based epidemiology. *Nature Water*, 1(5), 408-415. <https://doi.org/10.1038/s44221-023-00083-8>
- Sommer, J., Bobal, M., Bromberger, B., Mester, P. J., & Rossmanith, P. (2021). A new long-term sampling approach to viruses on surfaces. *Scientific Reports*, 11(1), 1-7. <https://doi.org/10.1038/s41598-021-96873-9>
- Stoufer, S., Soorneedi, A. R., Kim, M., & Moore, M. D. (2024). Sample Processing and Concentration Methods for Viruses from Foods and the Environment Prior to Detection. *Annual Review of Food Science and Technology*, 15(1), 225-246. <https://doi.org/10.1146/annurev-food-072023-034431>
- Strubbia, S., Schaeffer, J., Besnard, A., Wacrenier, C., Le Mennec, C., Garry, P., ... Le Guyader, F. S. (2020). Metagenomic to evaluate norovirus genomic diversity in oysters: Impact on hexamer selection and targeted capture-based enrichment. *International Journal of Food Microbiology*, 323, 108588. <https://doi.org/10.1016/j.ijfoodmicro.2020.108588>

REFERÈNCIES

- Tanner, B. D., Brooks, J. P., Gerba, C. P., Haas, C. N., Josephson, K. L., & Pepper, I. L. (2008). Estimated Occupational Risk from Bioaerosols Generated during Land Application of Class B Biosolids. *Journal of Environmental Quality*, 37(6), 2311-2321.
<https://doi.org/10.2134/jeq2007.0193>
- Thorn, J., & Kerekes, E. (2001). Health effects among employees in sewage treatment plants: A literature survey. *American Journal of Industrial Medicine*, 40(2), 170-179. <https://doi.org/10.1002/ajim.1085>
- Tiwari, A., Adhikari, S., Kaya, D., Islam, M. A., Malla, B., Sherchan, S. P., ... Pitkänen, T. (2023). Monkeypox outbreak: Wastewater and environmental surveillance perspective. *Science of the Total Environment*, 856(July 2022).
<https://doi.org/10.1016/j.scitotenv.2022.159166>
- U.S. EPA, United States Environmental Protection Agency (2005). Occurrence and Exposure Assessment for the Final Long Term 2 Enhanced Surface Water Treatment Rule. *EPA 815-R-06-002, EPA 815-R-06-002*.
- Ulrich, R. G., Drewes, S., Haring, V., Panajotov, J., Pfeffer, M., Rubbenstroth, D., ... Böhmer, M. M. (2023). Viral zoonoses in Germany: a One Health perspective. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 66(6), 599-616. <https://doi.org/10.1007/s00103-023-03709-0>
- Van der Fels-Klerx, H. J., Van Asselt, E. D., Raley, M., Poulsen, M., Korsgaard, H., Bredsdorff, L., ... Frewer, L. J. (2018). Critical review of methods for risk ranking of food-related hazards, based on risks for human health. *Critical Reviews in Food Science and Nutrition*, 58(2), 178-193.
<https://doi.org/10.1080/10408398.2016.1141165>
- Van Der Hoek, L. (2007). Human coronaviruses: What do they cause? *Antiviral Therapy*, 12(4 B), 651-658. <https://doi.org/10.1177/135965350701200s01.1>

- Vilsker, M., Moosa, Y., Nooij, S., Fonseca, V., Ghysens, Y., Dumon, K., ... de Oliveira, T. (2019). Genome Detective: an automated system for virus identification from high-throughput sequencing data. *Bioinformatics*, 35(5), 871-873. <https://doi.org/10.1093/bioinformatics/bty695>
- Wang, C. C., Prather, K. A., Sznitman, J., Jimenez, J. L., Lakdawala, S. S., Tufekci, Z., & Marr, L. C. (2021). Airborne transmission of respiratory viruses. *Science*, 373(6558). <https://doi.org/10.1126/science.abd9149>
- Wang, W., Kang, S., Zhou, W., & Vikesland, P. J. (2023). Environmental routes of virus transmission and the application of nanomaterial-based sensors for virus detection. *Environmental Science: Nano*, 10(2), 393-423. <https://doi.org/10.1039/d2en00600f>
- Wei, J., Jin, Y., Sims, T., & Kniel, K. E. (2011). Internalization of Murine Norovirus 1 by *Lactuca sativa* during Irrigation. *Applied and environmental microbiology*, 77, 2508-2512. <https://doi.org/10.1128/AEM.02701-10>
- WHO, World Health Organization (2011). *Guidelines for drinking-water quality: fourth edition incorporating the first addendum* [en línia]. Disponible a: <https://www.who.int/publications/i/item/9789241549950> [Últim accés: 09/04/2024].
- WHO, World Health Organization (2016). *Quantitative Microbial Risk Assessment: Application for Water Safety Management* [en línia]. Disponible a: <https://www.who.int/publications/i/item/9789241565370> [Últim accés: 09/04/2024].
- Wild, C. P. (2005). Complementing the genome with an «exposome»: The outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiology Biomarkers and Prevention*, 14(8), 1847-1850. <https://doi.org/10.1158/1055-9965.EPI-05-0456>

REFERÈNCIES

- Wolfe, M. K., Duong, D., Bakker, K. M., Ammerman, M., Mortenson, L., Hughes, B., ... Wigginton, K. R. (2022). Wastewater-Based Detection of Two Influenza Outbreaks. *Environmental Science and Technology Letters*, 9(8), 687-692. <https://doi.org/10.1021/acs.estlett.2c00350>
- Yao, L., Zhu, W., Shi, J., Xu, T., Qu, G., Zhou, W., ... Jiang, G. (2021). Detection of coronavirus in environmental surveillance and risk monitoring for pandemic control. *Chemical Society Reviews*, 50(6), 3656-3676. <https://doi.org/10.1039/d0cs00595a>
- Zambrana, W., & Boehm, A. B. (2023). Occurrence of Human Viruses on Fomites in the Environment: A Systematic Review and Meta-analysis. *ACS Environmental Au*, 3(5), 277-294. <https://doi.org/10.1021/acsenvironau.3c00025>
- Zaneti, R. N., Girardi, V., Spilki, F. R., Mena, K., Westphalen, A. P. C., da Costa Colares, E. R.,... Etchepare, R. G. (2021). Quantitative microbial risk assessment of SARS-CoV-2 for workers in wastewater treatment plants. *Science of the Total Environment*, 754, 142163. <https://doi.org/10.1016/j.scitotenv.2020.142163>
- Zaouri, N., Jumat, M. R., Cheema, T., & Hong, P. Y. (2020). Metagenomics-based evaluation of groundwater microbial profiles in response to treated wastewater discharge. *Environmental Research*, 180, 108835. <https://doi.org/10.1016/j.envres.2019.108835>
- Zhang, W., Li, L., Deng, X., Kapusinszky, B., & Delwart, E. (2014). What is for dinner? Viral metagenomics of US store bought beef, pork, and chicken. *Virology*, 468-470, 303-310. <https://doi.org/10.1016/j.virol.2014.08.025>

Annexos



Material suplementari de l'article 1

M&M

1. Wastewater samples and SARS-CoV-2 quantification

Urban 24-hours composite sewage samples were collected from February to May 2021 from different wastewater treatment plants (WWTP) located in Catalonia, northeast Spain. The biggest wastewater treatment facility, WWTP1, located at the city of Barcelona, serves half of the city's population and receives domestic and industrial waste from the sewer system. All WWTP are included in the Catalan Surveillance Network of SARS-CoV-2 in wastewater (www.sarsaigua.icra.cat), in which the authors of this study participate in measuring the SARS-CoV-2 RNA levels in wastewater samples collected weekly since July 2020. All samples were collected, transported at 4 °C to the laboratories (University of Barcelona, UB; Centre for Omic Sciences, COS) and processed the same day of arrival by ultrafiltration using CP-Select™ or Centricon® Plus-70 30 KDa, which have been previously described to be equivalent wastewater concentration methods for SARS-CoV-2 RNA detection (Forés et al., 2021). Concentration of viral particles was performed by first removing debris by centrifuging at 4750 x g for 30 min and processing 80 ml of the supernatant for concentration by ultrafiltration to obtain a final concentrate of 300 µl. All samples were spiked with the bacteriophage MS2 as a control before any processing was carried out. Nucleic acid extraction was done using QIAamp Viral RNA Mini Kit, using the QIAcube automatic system (Qiagen), recovered into a final volume of 70 µl. The SARS-CoV-2 concentration was calculated using specific RT-qPCR of the nucleocapsid protein (N1 and N2 assays) following the procedure described before (Rusiñol and Zammit et al., 2021). In particular, the 2019-nCoV RUO qPCR probe assay primer/probe mix (IDT) and the RNA Ultrasense™

One-Step RT-qPCR System (Invitrogen) were used to prepare the RT-qPCR mixture. Serial dilutions of Twist Synthetic SARS-CoV-2 RNA control 2, quantified using Qubit 3.0 dsDNA HS Assay Kit (Life Technologies), were used for standard curve construction. All RT-qPCR were performed in quadruplicate, with a 1/10 dilution of the nucleic acid extraction from each sample in order to characterize potential inhibitory effects, and included non-template controls.

2. Study of variants

Retrotranscription of nucleic acid extractions was performed using SuperScript IV enzyme (Life Technologies) and Random Hexamers (50 µM) following manufacturer instructions. Briefly, 11 µl of sample nucleic acid extraction was denaturalized after 5 min at 65 °C and 5 min on ice. After the addition of the SuperScript IV and RNase-OUT™ (Invitrogen), retrotranscription was performed over 10 min at 23 °C, 50 min at 50 °C and 10 min at 80 °C. The cDNA obtained was used as an input material for both approaches developed for the study of variants.

2.1 Nested PCR targeting signature mutations of the main VOCs and VOIs followed by Sanger sequencing

The cDNA obtained from the retrotranscription step was processed by specific nested PCR (nPCR) designed to capture signature mutations of different VOCs and VOIs. Most of the primers used were previously described as suitable for variant typing purposes (UG-HUG, 2020), but three new reverse nPCR primers were designed using Geneious software version 11.0 (<https://www.geneious.com>). These reverse nPCR primers were designed containing a signature mutation of a particular VOC within the primer sequence aimed to amplify a particular variant. Also, the primer

design considered comprising certain signature mutations located in the middle of the amplicon sequence. Using GISAID (<https://www.gisaid.org/>), primers were tested in silico against complete genomes belonging to the main VOCs and VOIs described to date: α , β , γ , δ , ϵ , ζ , η , θ , ι , κ , λ , (1st August 2021) by the European and the American Centre of Disease Control (ECDC, 2021 and CDC, 2021). A completed list with the primers used in this approach is showed in Table S1. The PCR products were cleaned and concentrated with Zymo DNA Clean & Concentrate kit (Zymo Research).

Table S1. List of primers used in nPCR assays targeting signature mutations of the main VOCs. Amplicons length expected are 529 or 538 bp (corresponding to the α variant or reference genome, respectively) for assay A, whereas amplicon length expected for assay B is 559 bp.

		Primer	Position	Sequence	Ref.
Assay A	First PCR	F44	21475-21498	TCTCTTCTTAGT AAAGGTAGACTT	(UG-HUG, 2020)
		R44	22044-22021	GAATAAACTCTG AACTCACTTTCC	
	nPCR	F44	21475-21498	TCTCTTCTTAGT AAAGGTAGACTT	(UG-HUG, 2020)
		DelY144	22012-21990	TTTGTTGTTTT GTGGTAAA	This study
Assay B	First PCR	F46	22476-22498	CCTTCACTGTAG AAAAAGGAATC	(UG-HUG, 2020)
		R46	23055-23032	TGGAAACCATAT GATTGTAAAGG A	
	nPCR	F46	22476-22498	CCTTCACTGTAG AAAAAGGAATC	(UG-HUG, 2020)
		E484K_R	23037-23015	GGAAAGTAACA ATTAAACCTTT	This study

All specific PCR and nPCR were performed using the DreamTaq™ DNA Polymerase (ThermoScientific). The amplification conditions for firsts rounds of PCR were as follows: 94 °C 5 min, followed by 35 cycles of 94 °C 1.30 min, 55 °C 1.30 min, 72 °C 1.30 min, and a final extension at 72 °C for 10 min. For nPCR reactions, the temperature of the annealing step was modified to 60 °C. All PCR reactions were performed using 1 µL of

each primer (25 μ M) and either 10 μ L of cDNA (first PCR) or 5 μ L of the first PCR product (nested PCR) in a final volume of 50 μ L.

The PCR products were cleaned and concentrated with Zymo DNA Clean & Concentrate kit (Zymo Research) and quantified using Qubit 3.0 dsDNA HS Assay Kit (Life Technologies). The purified DNA was directly sequenced by Sanger sequencing. All sequences were compared to the reference SARS-CoV-2 sequence NC_045512 and signature mutations were identified.

2.2 Amplicon Deep Sequencing of three different regions of S gene using Illumina MiSeq™

The cDNA obtained from the retrotranscription step was also processed by ADS using specific nPCR assays for three different regions of the spike gene (A1, A2 and A3), using primers previously described as suitable for variant typing purposes (UG-HUG, 2020). The nPCR primers were extended with the incorporation of Illumina® adapter at the 5' of the primers. A completed list with the primers used in this approach is shown in Table S2. The amplification conditions for PCR and nPCR reactions were as follows: 94 °C 5 min, followed by 35 cycles of 94 °C 1.30 min, 55 °C 1.30 min, 72 °C 1.30 min, and a final extension at 72 °C for 10 min.

Table S2. List of primers used in ADS assays targeting signature mutations of the main VOCs. Amplicons length expected are 570 bp for assay A1, 580 bp for assay A2 and 565 bp for assay A3. Primers extended with the incorporation of Illumina® adapters is marked as "-I" in the primer name.

		Primer name	Position	Sequence	Reference
Assay A1	First PCR	F44	21475-21498	TCTCTTCTTAGTAAAGGTAG ACTT	(UG-HUG, 2020)
		R45	22545-22523	CTAACAATAGATTCTGTTGG TTG	
	nPCR	F44-I	21475-21498	TCGTCGGCAGCGTCAGATG TGTATAAGAGACAGTCTCTT CTTAGTAAAGGTAGACTT	Modified from (UG-HUG, 2020)
		R44-I	22044-22021	GTCTCGTGGGCTCGGAGAT GTGTATAAGAGACAGGAATA AACTCTGAACTCACTTTCC	
Assay A2	First PCR	F46	22476-22498	CCTTCACTGTAGAAAAAGGA ATC	(UG-HUG, 2020)
		R47	23543-23520	CATATGAGTTGTTGACATGT TCAG	
	nPCR	F46-I	22476-22498	TCGTCGGCAGCGTCAGATG TGTATAAGAGACAGCCTTCA CTGTAGAAAAAGGAATC	Modified from (UG-HUG, 2020)
		R46-I	23055-23032	GTCTCGTGGGCTCGGAGAT GTGTATAAGAGACAGTGGAA ACCATATGATTGTAAAGGA	
Assay A3	First PCR	F46	22476-22498	CCTTCACTGTAGAAAAAGGA ATC	(UG-HUG, 2020)
		R47	23543-23520	CATATGAGTTGTTGACATGT TCAG	
	nPCR	F47-I	22979-22997	TCGTCGGCAGCGTCAGATG TGTATAAGAGACAGTATCAG GCCGGTAGCACAC	Modified from (UG-HUG, 2020)
		R47-I	23453-23520	GTCTCGTGGGCTCGGAGAT GTGTATAAGAGACAGCATAT GAGTTGTTGACATGTTTCAG	

Amplicons obtained were purified from agarose gel using QIAquick Gel Extraction (Qiagen) and sequenced, in parallel, with Illumina® MiSeq™ 2x300 bp platform.

Data analysis was performed by trimming primers from the reads with cut adapt (Martin, 2011) and then denoised into a set of amplicon sequence variants (ASVs) with DADA2 tool (Callahan et al., 2016). DADA2 was executed through QIIME2's DADA2 plugin, with the "denoise-single" command, with the following parameters: "--p-trunc-len 0 --p-trim-left 0 --

p-max-ee 2.0 --p-trunc-q 2 --p-pooling-method independent --p-chimera-method consensus --p-min-fold-parent-over-abundance 1.0 --p-n-reads-learn 1000000 --p-hashed-feature-ids True". The sequences of these ASVs were aligned against the SARS-CoV-2 Wuhan reference genome NC_045512.2 using MAFFT (Kato and Standley, 2013) with the following parameters: "--op 1.53 --ep 0.0 --maxiterate 0 --auto --thread -1".

With the ASV sequences aligned to the reference genome, the presence or absence of a specific set of signature mutations of the VOI and VOC for each sequence was checked, and that information was used to assess to which variants that sequence could belong to.

References:

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi:10.1038/nmeth.3869

Forés, E., Bo, S., Itarte, M., Martínez-puchol, S., Hundesa, A., Calvo, M., Borrego, C.M., 2021. Science of the Total Environment Evaluation of two rapid ultra filtration-based methods for SARS-CoV-2 concentration from wastewater 768. doi:10.1016/j.scitotenv.2020.144786

Kato, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* 30, 772–780. doi:10.1093/molbev/mst010

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10. doi:10.14806/ej.17.1.200

Rusiñol, M., Zammit, I., Itarte, M., Forés, E., Martínez-puchol, S., Girones, R., Borrego, C., 2021. Science of the Total Environment Monitoring waves of the COVID-19 pandemic: Inferences from WWTPs of different sizes 787. doi:10.1016/j.scitotenv.2021.147463

UG-HUG, 2020. Protocol for specific RT-PCRs for marker regions of the Spike region indicative of the UK SARS-CoV- 2 variant B.1.1.7 and the South African variant 501Y.V2.

ADS additional data

		Reads	ASV		
			Assay A1	Assay A2	Assay A3
WWTP 1 (1,497,767 inh.)	2/2/2021	4019460	108	129	34
	9/2/2021	2347220	108	130	28
	13/4/2021	631012	134	117	NT
	20/4/2021	428122	128	104	NT
	4/5/2021	247832	130	104	NT
WWTP 2 (183,517 inh.)	2/2/2021	2676466	111	125	30
	9/2/2021	2343200	107	122	20
	13/4/2021	336222	115	111	NT
	20/4/2021	93850	117	101	NT
	4/5/2021	301540	119	112	NT
WWTP 3 (68,860 inh.)	2/2/2021	1732266	100	121	25
	9/2/2021	780216	96	120	13
	13/4/2021	492588	128	112	NT
	20/4/2021	217250	116	108	NT
	4/5/2021	333824	126	106	NT

NT: not tested

Material suplementari de l'article 2

- Supplementary File 1. Complete SARS-CoV-2 assembled scaffold from wastewater contigs (*online*).

- Supplementary File 2. Complete list of all the identified viral species and families in the March wastewater sample by Target Enrichment NGS.

Family	Assignment	# Contigs
Adenoviridae	Human mastadenovirus A	1
	Human mastadenovirus F	13
	Simian mastadenovirus F	5
Alphaflexiviridae	Pepino mosaic virus	1
Anelloviridae	Avian gyrovirus 2	2
	Chicken anemia virus	1
Astroviridae	Astrovirus MLB1	2
	Astrovirus MLB2	2
	Astrovirus VA3	3
	Canine astrovirus	1
	Canine astrovirus	2
	Canine astrovirus	1
	Feline astrovirus 2	1
	Feline astrovirus 2	1
	Feline astrovirus 2	1
	Feline astrovirus 2	1
	Feline astrovirus 2	1
	Feline astrovirus 2	1
	Feline astrovirus 2	1
	HMO Astrovirus A	3
	Mamastrovirus 1	1
	Mamastrovirus 2	1
	Mamastrovirus 9	2
	Porcine astrovirus 4	2

Family	Assignment	# Contigs
Betaflexiviridae	Garlic common latent virus	6
Caliciviridae	Norovirus GI	3
	Norovirus GII	2
	Norway rat hunnivirus	4
	Sapporo virus	5
	Sapporo virus	1
Circoviridae	Artemia melana sponge associated circular genome	1
	Cyclovirus TN25	2
	Human associated cyclovirus 10	1
	Human circovirus VS6600022	2
	Human circovirus VS6600022	2
	Human fecal virus Jorvi2	2
	Porcine circovirus 2	1
	Porcine stool-associated circular virus 4	2
Coronaviridae	Betacoronavirus 1 (HCoV OC43)	1
	Feline coronavirus	1
	Lucheng Rn rat coronavirus	5
	SARS-CoV-2	8
	Canine coronavirus	1
Dicistroviridae	Aphid lethal paralysis virus	4
	Drosophila C virus	5
	Rhopalosiphum padi virus	1
Genomoviridae	Gopherus associated genomovirus 1	2
	Human associated gemykibivirus 2	1
	Rhinolophus associated gemykibivirus 1	1
	Sewage derived gemycircularvirus 2	3
	Sewage derived gemycircularvirus 3	2
	Sewage-associated gemycircularvirus-10a	1
Hepeviridae	Hepatitis E virus	4
	Hepatitis E virus rat/R63/DEU/2009	4
Iflaviridae	Culex Iflavi-like virus 1	2
	Sacbrood virus	3
	Varroa destructor virus 1	4
	Yongsan iflavirus 1	1

Family	Assignment	# Contigs
Inoviridae	uncultured phage WW-nAnB	3
	Uncultured phage WW-nAnB strain 2	4
	Uncultured phage WW-nAnB strain 3	5
Iridoviridae	Invertebrate iridescent virus 31	2
Leviviridae	Acinetobacter phage AP205	1
Marnaviridae	Sanfarnavirus 2	4
Microviridae	Chimpanzee faeces associated microphage 1	1
	Escherichia virus G4	2
	Escherichia virus ID52	2
	Escherichia virus Talmos	1
	Gokushovirinae Bog1183_53	3
Mimiviridae	Acanthamoeba polyphaga mimivirus	1
Myoviridae	Acinetobacter phage Acj61	4
Nodaviridae	Le Blanc nodavirus (segment RA11)	1
	Orsay virus (segment RNA 1)	1
	Orsay virus (segment RNA 2)	1
	Santeuil nodavirus (segment RNA 2)	1
Paramyxoviridae	Human respirovirus 3	1
Partitiviridae;	Rosellinia necatrix partitivirus 8 (segment RNA 1)	1
Partiviridae	Fusarium poae partitivirus 2 (segment 2)	1
Parvoviridae	Adeno-associated virus - 1	2
	Adeno-associated virus - 2	1
	Adeno-associated virus - 5	1
	Aedes albopictus densovirus	1
	Avian adeno-associated virus strain DA-1	2
	Blattodean ambidensovirus 1	2
	Blattodean ambidensovirus 2	3
	Bovine adeno-associated virus	1
	Carnivore bocaparvovirus 1	4
	Carnivore bocaparvovirus 2	3
	Carnivore protoparvovirus 1	4
	Decapod penstyldensovirus 1	2
	Dipteran ambidensovirus 1	1
	Human bocavirus 3	2
	Mythimna loreyi densovirus	3
...		

Family	Assignment	# Contigs
... Parvoviridae	Orthopteran ambidensovirus 1	3
	Parus major densovirus	2
	Porcine bocavirus 3	1
	Porcine parvovirus	2
	Primate bocaparvovirus 1	1
	Primate bocaparvovirus 2	2
	Primate protoparvovirus 1	3
	Rodent protoparvovirus 1	2
	Rodent protoparvovirus 2	4
	Ungulate bocaparvovirus 3	2
	Ungulate bocaparvovirus 4	3
Picobirnaviridae	Chicken picobirnavirus (segment RNA 2)	1
	Human picobirnavirus (segment 1)	1
	Human picobirnavirus (segment 2)	1
	Human picobirnavirus (segment 2)	2
	Otarine picobirnavirus (segment 2)	1
	Picobirnavirus dog/KNA/2015	2
	Picobirnavirus green monkey/KNA/2015	3
	Porcine picobirnavirus (segment S)	1
	Porcine picobirnavirus (segment S)	2
Picornaviridae	Aichivirus A	2
	Aichivirus F	2
	Cardiovirus C	4
	Chicken megrivirus	1
	Chicken picornavirus 1	2
	Chicken picornavirus 5	2
	Cosavirus A	4
	Cosavirus D	1
	Enterovirus A	3
	Enterovirus B	5
	Enterovirus C	5
	Parechovirus A	2
	Rabovirus	4
	Rhinovirus A	2
	Rhinovirus B	2
...		

Family	Assignment	# Contigs
... Picornaviridae	Salivirus A	3
	Sicinivirus A	4
	Theilovirus	1
	Theilovirus	3
Polyomaviridae	Human polyomavirus 1	2
	Human polyomavirus 2	1
Reoviridae	Mammalian orthoreovirus (segment L3)	2
Satellites	Sewage-associated circular DNA molecule-1	1
	Sewage-associated circular DNA molecule-3	1
Secoviridae	Broad bean wilt virus 2 (segment RNA 1)	1
	Broad bean wilt virus 2 (segment RNA 1)	2
	Broad bean wilt virus 2 (segment RNA 1)	1
	Broad bean wilt virus 2 (segment RNA 2)	1
	Carrot torradovirus 1 (segment RNA 2)	2
	Squash mosaic virus (segment RNA 1)	4
	Squash mosaic virus (segment RNA 2)	2
Siphoviridae	Streptococcus phage Str-PAP-1	1
Smacoviridae	Human associated porprismacovirus 2	1
Solemoviridae	Papaya lethal yellowing virus	2
Tombusviridae	Grapevine Algerian latent virus	2
	Moroccan pepper virus	1
	Pelargonium leaf curl virus	2
Tymoviridae	Poinsettia mosaic virus	1
	Turnip yellow mosaic virus	3
Unclassified	Gyrovirus 4	2
	Hubei picorna-like virus 15	1
	Hubei picorna-like virus 61	5
	Hubei picorna-like virus 62	4
	Hubei picorna-like virus 63	4
	Hudisavirus sp.	3
	Hudisavirus sp.	1
	Human feces pecovirus	1
	Odonata-associated circular virus-18	1
	... Picalivirus A	3

Family	Assignment	# Contigs
... Unclassified	Picornavirales Tottori-HG1	4
	Wenzhou picorna-like virus 47	3
	Wuhan insect virus 23 (segment Seg 2)	1
Virgaviridae	Bell pepper mottle virus	4
	Cucumber green mottle mosaic virus	3
	Paprika mild mottle virus	1
	Pepper mild mottle virus	1
	Tobacco mild green mosaic virus	5
	Tobacco mosaic virus	2
	Tomato brown rugose fruit virus	1
	Tomato mosaic virus	1
	Turnip vein-clearing virus	1

- **Supplementary File 3.** Summary of sequencing results for the three samples considered on this work.

Sample	Wastewater	Clinical_A	Clinical_B
Raw PE Reads (x2)	5,901,370	6,021,795	4,663,999
Raw PE Total bp	1,850,065,979	3,015,620,573	2,353,498,179
Clean PE Reads (x2)	5,006,516	375,769	316,398
Clean PE Total bp	1,399,002,141	74,651,907	66,525,983
Clean SG Reads	823,018	5,535,930	4,238,097
Clean SG Total bp	139,990,580	441,056,111	337,855,561

- **Supplementary File 4.** Mafft whole alignment from the recovered reads with respect to the alignment with 1200 randomly chosen GISAID SARS-CoV-2 sequences (*online*).

Material suplementari de l'article 3

- Supplementary information 1. Sequences used for the design of the probe panel applied in this study.

(A continuació)

Supplementary information 1. Sequences used for the design of the probe panel applied in this study.

Specie	Host	Sequence ID	Region	Size (bp)	Sequence description (complete genomes)
SARS-CoV-2	Human	NC_045512.2	Complete Genome	29903	Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1
SARS-CoV-2	Human	MW969753.1	Complete Genome	29819	Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/IND/CD211295/2021
SARS-CoV	Human	NC_004718	Complete Genome	29751	SARS coronavirus Tor2, complete genome
MERS-CoV	Human	NC_019843.3	Complete Genome	30119	Middle East respiratory syndrome-related coronavirus isolate HCoV-EMC/2012, complete genome
HCoV-OC43	Human	NC_006213.1	Complete Genome	30741	Human coronavirus OC43 strain ATCC VR-759
HCoV-229E	Human	NC_002645.1	Complete Genome	27317	Human coronavirus 229E
HCoV-NL63	Human	NC_005831.2	Complete Genome	27553	Human Coronavirus NL63
HCoV-HKU1	Human	NC_006577.2	Complete Genome	29926	Human coronavirus HKU1
PorCoV-HKU15	Porcine	NC_039208.1	Complete Genome	25425	Porcine coronavirus HKU15 strain HKU15-155
BCoV-ENT	Bovine	NC_003045.1	Complete Genome	31028	Bovine coronavirus isolate BCoV-ENT
MHV-S	Murine	NC_048217	Complete Genome	31335	Murine hepatitis virus strain A59
CCoV	Canine	NC_038861.1	Complete Genome	28586	Transmissible gastroenteritis virus, genomic RNA
FCoV	Feline	NC_002306.3	Complete Genome	29355	Feline infectious peritonitis virus
Avian infectious bronchitis virus	Avian	NC_001451.1	Complete Genome	27608	Avian infectious bronchitis virus
Alphacoronavirus Bat-CoV	Bat	NC_046964	Complete Genome	28128	Alphacoronavirus Bat-CoV/P.kuhlii/Italy/3398-19/2015
Bat-CoV 1A	Bat	NC_010437	Complete Genome	28326	Bat coronavirus 1A
Bat-CoV HKU10	Bat	NC_018871	Complete Genome	28494	Rousettus bat coronavirus HKU10
Bat-CoV HKU8	Bat	NC_010438	Complete Genome	28773	Bat coronavirus HKU8
Bat-CoV HKU2	Bat	NC_009988	Complete Genome	27165	Bat coronavirus HKU2

Material suplementari de l'article 4

- **Table S1.** Quantification of viral pathogens in irrigation water and fresh produce samples. Values are expressed in GC/L for irrigation water samples and in GC/25 g for food samples. ND: Not Detected.

(A continuació)

- **Table S2.** Reads obtained from each viral assignment in irrigation water, strawberry, lettuce and parsley samples using target enrichment sequencing. Colors represent the specific host of vertebrate viruses, legend can be found under the table.

(A continuació)

Supplementary Files

Table S1. Quantification of viral pathogens in irrigation water and fresh produce samples. Values are expressed in GC/L for irrigation water samples and in GC/25 g for food samples. ND: Not Detected.

		HAdV	NoV GI	NoV GII	HEV
Production Site 1	GW1.1	ND	ND	2.37×10^2	ND
	GW1.2	ND	ND	2.51×10^2	ND
	GW1.3	ND	ND	ND	ND
	GW1.4	ND	ND	ND	ND
	LET1.1	ND	ND	ND	ND
	LET1.2	ND	ND	ND	ND
	LET1.3	ND	ND	ND	ND
	LET1.4	ND	ND	4.35×10^2	ND
	STR1.1	ND	ND	3.26×10^2	ND
	STR1.2	ND	ND	2.55×10^2	ND
	STR1.3	ND	ND	ND	ND
	STR1.4	ND	ND	ND	ND
	STR1.5	ND	ND	ND	ND
	STR1.6	ND	ND	ND	ND
	STR1.7	ND	ND	ND	ND
	STR1.8	ND	ND	ND	ND
	PAR1.1	6.05×10^2	ND	ND	ND
	PAR1.2	ND	ND	ND	ND
	PAR1.3	ND	ND	ND	ND
	PAR1.4	ND	ND	ND	ND
Production Site 2	GW2.1	ND	ND	9.02×10^0	ND
	GW2.2	ND	ND	1.96×10^2	ND
	GW2.3	ND	ND	ND	ND
	GW2.4	ND	ND	ND	ND
	LET2.1	1.40×10^1	ND	ND	ND
	LET2.2	ND	ND	ND	ND
	LET2.3	3.02×10^2	ND	ND	ND
	LET2.4	ND	ND	3.35×10^2	ND
Production Site 3	RIV3.1	ND	ND	1.34×10^1	ND
	RIV3.2	ND	ND	ND	ND
	RIV3.3	ND	ND	ND	ND
	RIV3.4	ND	ND	ND	ND
	LET3.1	1.03×10^2	ND	ND	ND
	LET3.2	ND	ND	ND	ND
	LET3.3	ND	ND	ND	ND
	LET3.4	1.10×10^2	ND	ND	ND
	STR3.1	ND	ND	ND	ND
	STR3.2	4.16×10^1	ND	5.60×10^2	ND
	STR3.3	5.27×10^1	ND	ND	ND
	STR3.4	ND	ND	ND	ND
	PAR3.1	ND	ND	2.44×10^1	ND
	PAR3.2	8.83×10^2	ND	7.12×10^1	ND
	PAR3.3	8.76×10^2	ND	9.68×10^0	ND
	PAR3.4	ND	ND	9.64×10^0	ND
Llobregat river	LLRIV.1	3.47×10^4	6.36×10^3	1.33×10^4	ND
	LLRIV.2	5.99×10^3	1.01×10^4	4.75×10^4	ND
Treated wastewater	SE	8.99×10^4	3.77×10^3	ND	ND
	WFPE	1.88×10^5	9.49×10^3	6.64×10^2	ND

Table S2. Reads obtained from each viral assignment in irrigation water, strawberry, lettuce and parsley samples using target enrichment sequencing. Colors represent the specific host of vertebrate viruses, legend can be found under the table.

Host	Family	Assignment	Production Site 1		Production Site 2	Production Site 3				Llobregat river	Treated wastewater	
			GW1	STR1.1	GW2	RIV3	LET3.4	PAR3.2	LLRIV.1	SE	WFPE	
Bacteria	<i>Inoviridae</i>	Uncultured phage WW-nAnB			12				49			
	<i>Leviviridae</i>	Acinetobacter phage AP205				3						
		Escherichia virus MS2			90	62		4	6			18,808
		Bdellovibrio virus MH2K							10			
		Escherichia virus phiX174					4	2	4			
	<i>Microviridae</i>	Gokushovirinae Bog1183_53							2			
		Gokushovirinae GAIR4	18,239	18	38				53			22
		Gokushovirinae GNX3R			7							
	<i>Myoviridae</i>	Salmonella phage RE-2010		4								
		Salmonella virus Fels2		16								
	<i>Podoviridae</i>	Uncultured crAssphage										2
		Salicola phage CG-phi29				14						
Human	<i>Siphoviridae</i>	Salmonella virus SP31			61							
		Serratia phage Eta			385							
		Streptococcus virus Sfi19							20			
	<i>Astroviridae</i>	Astrovirus MLB1							4			
		Astrovirus MLB2							6			
		HMO Astrovirus A							18			
		Mamastrovirus 1				255			752			24
	<i>Caliciviridae</i>	Norovirus GI							20			
		Norovirus GII						3	115			
	<i>Circoviridae</i>	Human associated cyclovirus 6	2,026							18		

	<i>Parvoviridae</i>	Adeno-associated virus - 2			3	187			573		89
		Human bocavirus 3							62		
	<i>Picornaviridae</i>	Aichivirus A				73			155		
		Salivirus A							40	101	90
Invertebrates		Aphid lethal paralysis virus	2,013	2	8	135				3,798	
		Big Sioux River virus	18								18
	<i>Dicistroviridae</i>	Kashmir bee virus	437		10						
		Rhopalosiphum padi virus		4						11,283	
	<i>Genomoviridae</i>	Tick associated genomovirus 3	8,254								
		Sacbrood virus	24			14					
	<i>Iflaviridae</i>	Varroa destructor virus 1	28								
		Invertebrate iridescent virus 31				4					
	<i>Iridoviridae</i>	Anopheles gambiae densovirus							6		
	<i>Parvoviridae</i>	Dipteran ambidensovirus 1							28		
		Junonia coenia densovirus				4					
	<i>Unclassified</i>	Giant house spider associated circular virus 2			190						
		Odonata-associated circular virus-18									18
Other Vertebrates	<i>Unclassified Picornavirales</i>	Biomphalaria virus 2							2		
		Beihai permutotetra-like virus 2	2								
	<i>Unclassified Riboviria</i>	Changjiang picorna-like virus 14				10					
		Hubei picorna-like virus 15				8				28	
		Hubei picorna-like virus 51				8					
	<i>Astroviridae</i>	Feline astrovirus 2				139			586		
		Porcine astrovirus 4				54					
	<i>Caliciviridae</i>	Sapovirus GIII							16		
		Equine associated gemycircularvirus 1	66								
	<i>Genomoviridae</i>	Gerygone associated gemycircularvirus 1			4						
		Gerygone associated gemycircularvirus 2	503								
		Pteropus associated gemycircularvirus 2			73						
		Sewage derived gemykbivirus 1				24					

		Carnivore bocaparvovirus 2									152		21
Parvoviridae		Parus major densovirus	24										
		Porcine bocavirus 3									12		
		Porcine bocavirus 4-1									50		
		Porcine parvovirus									6		
		Primate bocaparvovirus 2									65		8
		Rodent protoparvovirus 1									76		
	Ungulate bocaparvovirus 4									7			
Picornaviridae		Norway rat humnivirus									8		
Alphaflexviridae		Pepino mosaic virus									10		
Rhabdoviridae		Lettuce big-vein associated varicosavirus					14						
		Cucumber necrosis virus									12		
Tombusviridae		Moroccan pepper virus				6							
		Tobacco necrosis virus D						4					
		Tomato bushy stunt virus				4							
		Tobacco virotopvirus 1						11					
Unclassified Riboviria		Bell pepper mottle virus									4		
Virgaviridae		Cucumber green mottle mosaic virus						153			27	14	13
		Pepper mild mottle virus						138			1,404	172	44
		Tobacco mosaic virus									4		
		Tomato mosaic virus						52	2		3,082	984	266
Genomoviridae		Sewage derived gemycircularvirus 1				80							
		Sewage derived gemykibivirus2				6							
		Sewage-associated circular DNA virus-20											
		Sewage-associated circular DNA virus-27						98				23	



Material suplementari de l'article 5

- **Table S1.** Viral assignments obtained in winter highly urbanized groundwater using target enrichment sequencing.

(A continuació)

- **Table S2.** Viral assignments obtained in winter industrial district groundwater using target enrichment sequencing.

(A continuació)

- **Table S3.** Viral assignments obtained in winter riverside groundwater using target enrichment sequencing.

(A continuació)

- **Table S4.** Viral assignments obtained in medium urbanized runoff sample RO1 using target enrichment sequencing.

(A continuació)

- **Table S5.** Viral assignments obtained in medium urbanized runoff sample RO2 using target enrichment sequencing.

(A continuació)

Table S1. Viral assignments obtained in winter highly urbanized groundwater using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average lenght (bp)	Host
Circoviridae	Circovirus	Bat circovirus	-	1329650	-	92	1	88,3	392 Vertebrate
		Pigeon circovirus	-	1414603	-	22	-	98,9	142 Vertebrate
Dicistroviridae	Cripavirus	Cripavirus mortiferum	Aphid lethal paralysis virus (ALPV)	209529	-	94	-	98,4	138,8 Invertebrate
		Cripavirus rapadi	Rhopalosiphum padi virus (RHPV)	66834	-	16	-	100	123 Invertebrate
Genomoviridae	Unclassified Genomoviridae	Eucommia ulmoides genomoviridae	-	2739878	-	20	-	92,8	127,7 Plant
Iflaviridae	Iflavirus	Iflavirus spexiguae	Spodoptera exigua iflavirus 1 (SelV-1)	1111710	-	64	-	92,3	122,2 Invertebrate
Parvoviridae	Iteradensovirus	Danaus plexippus plexippus iteravirus	-	1469956	-	24	-	97,7	129 Invertebrate
Picornaviridae	Kobuvirus	Aichivirus A	Canine kobuvirus (CaKov)	72149	-	360	1	95,8	755 Vertebrate
	Megivirus	Megivirus B	Pigeon mesivirus 2 (MeV-B2)	1379161	-	90	-	97	138,9 Vertebrate
Virgaviridae	Tobamovirus	Tobacco mosaic virus	-	12242	-	200	1	97,3	152,2 Plant

Table S2. Viral assignments obtained in winter industrial district groundwater using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average lenght (bp)	Host
Diatroviridae	Circovirus	Bat circovirus	-	1329650	3942	1	84,4	795	Vertebrate
	Cripavirus	Cripavirus mortiferum	Aphid lethal paralysis virus (ALPV)	209529	68	-	99,8	108,5	Invertebrate
	Unclassified Microviridae	Microviridae sp.	-	2202644	1376	1	84,6	162,7	-
Papillomaviridae	Betapapillomavirus	Betapapillomavirus 2	HPV-38 (Human papillomavirus 38)	333924	56	-	98,3	127,3	Human

Table S3. Viral assignments obtained in winter riverside groundwater using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average length (bp)	Host
Circoviridae	Circovirus	Pigeon circovirus	-	1414603	58	1	93,1	247	Vertebrate
		Sewage associated circovirus	-	689405	160	1	85,1	268	-
Dicistroviridae	Cripavirus	Cripavirus mortiferum	Aphid lethal paralysis virus (ALPV)	209529	162	2	99,5	201	Invertebrate
Genomoviridae	Unclassified Genomoviridae	Genomoviridae sp.	-	2202565	228	1	89,4	143,3	-
		Plant associated genomovirus 6	-	2584400	40	1	97,4	306	Plant
Iflaviridae	Ifavirus	Ifavirus aladeformis	Deformed wing virus (DWV)	198112	319	2	98,5	182,4	Invertebrate
		Varrcoa destructor virus 1	-	232800	87	-	99,3	239	Invertebrate
Microviridae	Unclassified Microviridae	Blackfly microvirus SF02	-	2576452	203	1	83,8	131,4	Bacteria
		Gokushovirinae environmental samples	-	1478972	340	1	83,6	144,2	-
Papillomaviridae	Betapapillomavirus	Betapapillomavirus 2	HPV-38 (Human papillomavirus 38)	333924	22	-	99	151,5	Human
Parvoviridae	Unclassified Parvovirinae	Parvovirus NIH-CQV*	-	1341019	90	1	87,5	319	Human

* While a study associated Parvovirus NIH-CQV with patients having seronegative hepatitis by NGS (1), it has been reported that these sequences might be laboratory-derived contaminants from nucleic acid extraction spin columns (2).

1. Xu B, Zhi N, Hu G, Wan Z, Kajigaya S, Zhao K, et al. Identification and Characterization of a Novel Parvovirus-Like Virus in Seronegative Hepatitis Patients by Next Generation Sequencing. Blood. 2012;120(21):273–273.

2. Naccache SN, Hackett J, Delwart EL, Chiu CY. Concerns over the origin of NIH-CQV, a novel virus discovered in Chinese patients with seronegative hepatitis. Proc Natl Acad Sci U S A. 2014;111(11):2014.

Table S4. Viral assignments obtained in medium urbanized runoff sample RO1 using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average length (bp)	Host
Anelloviridae	Gyrovirus	Gyrovirus chickenanemia	CAV (Chicken anemia virus)	12618	228	1	99.8	495	Bird
		Avastrovirus 2	-	1239438	80	1	89.5	333	Human/Vertebrate
Astroviridae	Unclassified Astroviridae	Avian astrovirus	-	1914440	156	1	91	184.7	Vertebrate
		Feral pigeon astrovirus	-	928291	2902	3	93.5	475.1	Vertebrate
		Wood pigeon astrovirus	-	928292	68	1	91.9	236	Vertebrate
Autographiviridae	Caracivirus	Citrobacter virus CR8	-	2733611	30	-	100	150.5	Bacteria
		Shingomonas virus Scott	-	2733954	166	2	92.1	153.3	Bacteria
Caliciviridae	Vesivirus	Calicivirus isolate Geel 2008/Belgium	-	673765	1917	2	91.2	1132	Vertebrate
		Canine vesivirus	-	1176484	2698	5	94.7	812.7	Vertebrate
Circoviridae	Circovirus	Vesivirus 2117	-	241631	29	-	93.1	128	Vertebrate
		Circovirus canine	-	1194757	44	1	97.5	325	Vertebrate
		Circovirus starling	-	349370	26	-	97.3	110	Vertebrate
		Pigeon circovirus	-	1414603	9412	5	98.1	761.9	Vertebrate
		Cyclovirus bashi	Human associated cyclovirus 4 (HuCYV4)	2038722	156	-	99	109.6	Human
	Cyclovirus	Cyclovirus femage	Bat associated cyclovirus 2 (BaCYV-2)	2050585	175	1	84.8	256	Vertebrate
		Cyclovirus kriptero	Bat associated cyclovirus 3 (BaCYV-3)	2050586	636	1	82.5	235.7	Vertebrate
		Cyclovirus TN10	-	742955	82	1	99.6	233	Human
		Cyclovirus TN12	-	742953	605	1	99.9	214.3	Human
		Alphacoronavirus 1	Feline coronavirus (FECV)	693997	42	-	93.1	140.5	Vertebrate
Coronaviridae	Alphacoronavirus	Apavavirus apisacutum	Acute bee paralysis virus (ABPV)	92444	511	3	97.7	263.5	Invertebrate
		Apavavirus kashmirens	Kashmir bee virus (KBV)	68876	686	5	97	357.2	Invertebrate
		Apavavirus taurans	Taura syndrome virus (TSV)	142102	414	1	93	875	Invertebrate
	Cripavirus	Aphis gossypii virus	-	2448899	948	3	98.7	400.3	Invertebrate
		Cripavirus mortiferum	Aphid lethal paralysis virus (ALPV)	209529	6109	11	97.7	1273.9	Invertebrate
Dicistroviridae	Unclassified Dicistroviridae	Cripavirus ropadi	Rhopalosiphum padi virus (RHPV)	66834	14608	9	98.6	3522.8	Invertebrate
		Macrobacchium rosenbergii Taihu virus	-	1027947	40	1	86.6	246	Invertebrate
Dreileviridae	Souteliovirus	Pantoea virus AAS23	-	2734030	66	-	99.1	145.5	Bacteria
		Common bean-associated gemycircularvirus	-	1897538	15	-	84.8	112	Plant
Gemoviridae	Gemycircularvirus	Gemycircularvirus porc1	Porcine associated gemycircularvirus 1 (PorAGcV-1)	1985393	70	-	90.6	140.1	Vertebrate
		Giant panda associated gemycircularvirus	-	2016461	232	1	85.8	189.4	Vertebrate
		Gila monster-associated gemycircularvirus	-	2249931	62	1	97.6	288	Vertebrate
		Gemycovirus pteroi	Pteropus associated gemycovirus 1 (PrGgV-1)	1985427	126	-	89.8	113.4	Vertebrate
		Gemycovirus sewapoi	Sewage derived gemycovirus 1 (SDGgV-1)	1985428	2257	1	88.9	608.7	Human
	Gemycibivirus	Gemycibivirus confami	Canine feces-associated gemycircularvirus (CaFAGcV)	2017709	34	1	93.8	215.6	Vertebrate
		Gemycibivirus haeme3	Finch associated gemycovirus 2 (FaGgV-2)	2576454	94	1	94.5	197.2	Vertebrate
		Gemycibivirus humas3	Human associated gemycibivirus 5 (HAGKIV-5)	2004962	1179	5	96.5	392.8	Human
		Gemycolovirus deraoi	Tick-associated genomovirus 2 (TaGgV-2)	2025478	6234	1	87	808	Invertebrate
		Gemycovangvirus humas1	Human associated gemycovangvirus 1 (HAGTV-1)	1985415	34	-	99	150.5	Human
Genomoviridae	Unclassified Genomoviridae	Agrotis ipsilon-associated genomovirus	-	2656916	1164	1	80.9	1052.2	Invertebrate
		Apis mellifera genomovirus 1	-	2563837	32	-	90.1	126	Invertebrate
		Plant associated genomovirus 8	-	2584402	292	-	95	126.4	Plant
		Tortoise genomovirus 3	-	2582886	330	2	98.4	418.8	Vertebrate

<i>Iflaviridae</i>	<i>Iflavirus</i>	<i>Iflavirus bladeformis</i> <i>Iflavirus spexiguae</i> <i>Varroa destructor virus 1</i>	Deformed wing virus (DWV) <i>Spodoptera exigua iflavirus 1</i> (SeV-1)	198112 1111710 4761 232800 346 1654357 16	134361 4761 346 16	12 5 - - - - -	98.8 89.4 97 98.5	7284.8 Invertebrate 690.5 Invertebrate 123.35 Invertebrate 101 Invertebrate
<i>Nodaviridae</i>	Unclassified <i>Iflaviridae</i>	<i>La Jolla virus</i>	-	12287 1423 12288 406	1423 406	2 3	84.3 87.2	504.3 Invertebrate 261 Invertebrate/Vertebrate
	<i>Alphanodavirus</i>	<i>Nodamura virus</i>	-	1108949 160	160	-	94.7	127.2 Vertebrate
	<i>Bocaparvovirus</i>	<i>Canine bocavirus</i>	-	1511866 40	40	1	99.3	292 Vertebrate
		<i>Carnivore bocaparvovirus 1</i> (CPV-1)	Canine parvovirus 1 (CPV-1)	1511867 58	58	1	93.4	228 Vertebrate
	<i>Densovirus</i>	<i>Canine bocavirus 1</i> (CBov-1)	Canine bocavirus 1 (CBov-1)	2781641 20	20	1	87.8	196 Vertebrate
		<i>Blatella germanica densovirus-like virus 1</i>	-	2776787 18	18	-	100	151 Invertebrate
		<i>Adeno-associated virus</i>	-	272636 257	257	3	82.4	270.5 Human/Vertebrate
<i>Parvoviridae</i>	<i>Dependoparvovirus</i>	<i>Anseriform dependoparvovirus 1</i>	Goose parvovirus (GPV)	1511895 320	320	1	84.1	208.9 Vertebrate
		<i>Avian dependoparvovirus 1</i>	-	1511894 178	178	1	84	276.4 Vertebrate
		<i>Marsupial adeno-associated virus 1</i>	-	2497568 175	175	1	85.4	166.3 Vertebrate
	<i>Iteradensovirus</i>	<i>Hordeum marinum itera-like densovirus</i>	-	1561946 60	60	1	96.3	219 Plant
		<i>Lepidopteran iteradensovirus 2</i>	-	1513211 18	18	-	94.4	151 Invertebrate
	<i>Scindambidensovirus</i>	<i>Orthopteran scindambidensovirus 1</i>	Acheta domestica densovirus (AdDNV)	2745121 1010	1010	5	99.6	343.5 Invertebrate
	Unclassified <i>Parvoviridae</i>	<i>Pigeon parvovirus A</i>	-	1399967 11718	11718	10	94.8	998.9 Vertebrate
<i>Picobirnaviridae</i>	<i>Picobirnavirus</i>	<i>Picobirnavirus dog/XNA/2015</i>	-	1961162 76	76	-	98.5	150.8 Vertebrate
	<i>Kobuvirus</i>	<i>Aichivirus A</i>	Canine kobuvirus (Cakov)	1836608 635	635	1	95.45	334.1 Vertebrate
<i>Picornaviridae</i>	<i>Megrivirus</i>	<i>Megrivirus B</i>	Pigeon mesivirus 1 (MeV-B1)	2170005 821	821	4	90	234.2 Vertebrate
	<i>Mischivirus</i>	<i>Mischivirus D</i>	Pigeon mesivirus 2 (MeV-B2)	1379161 22712	22712	12	92	1572 Vertebrate
<i>Polycipiviridae</i>	<i>Sapolycivirus</i>	<i>Lasius niger virus 1</i>	Canine picornavirus (CanPV)	2560598 132	132	1	93.3	267 Vertebrate
<i>Retroviridae</i>	<i>Alpharetrovirus</i>	<i>Avian leukosis virus</i>	-	2018503 24	24	-	94	100 Invertebrate
		<i>Cucumber green mottle mosaic virus</i>	-	11864 204	204	1	99.8	405.7 Vertebrate
		<i>Hibiscus latent Fort Pierce virus</i>	-	12235 30	30	-	99.8	117.3 Plant
		<i>Hibiscus latent Singapore virus</i>	-	233051 546	546	6	99.5	216.1 Plant
<i>Virgaviridae</i>	<i>Tobamovirus</i>	<i>Pepper mild mottle virus</i>	-	185955 908	908	6	99.6	292.5 Plant
		<i>Rehmannia mosaic virus</i>	-	12239 28	28	-	99.3	151 Plant
		<i>Tobacco mild green mosaic virus</i>	-	425279 260	260	1	96.4	146.3 Plant
		<i>Tobacco mosaic virus</i>	-	12241 444	444	4	99.1	256.2 Plant
		<i>Tomato mosaic virus</i>	-	12242 11358	11358	18	97.8	886 Plant
<i>Tombusviridae</i>	<i>Tombusvirus</i>	<i>Cowpea tomosvirus 1</i>	-	12253 228	228	-	99.4	125.9 Plant
<i>Tymoviridae</i>	Unclassified <i>Tymoviridae</i>	<i>Poinsettia mosaic virus</i>	-	1913126 184	184	1	90.5	106.6 Plant
		-	-	113553 10622	10622	9	87.4	2306.3 Plant

Table S5. Viral assignments obtained in medium urbanized runoff sample RO2 using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average lenght (bp)	Host
Anelloviridae	Thetatorquevirus	Torque teno canid virus 1	Torque teno canis virus (TTCaV)	687385	24	-	99.2	128	Vertebrate
	Mamastrovirus	Canine astrovirus	-	1157338	21	-	97.4	141.7	Vertebrate
Circoviridae	Circovirus	Pigeon circovirus	-	1414603	100	1	98.1	187.4	Vertebrate
	Cyclovirus	Cyclovirus humana	Human associated cyclovirus 10 (HucV10)	2038728	41	-	93.4	104.5	Human
		Cyclovirus TN12	-	742953	3045	1	99	400.8	Human
Dicistroviridae	Aparavirus	Aparavirus apisacutum	Acute bee paralysis virus (ABPV)	92444	48	-	96.4	100	Invertebrate
	Cripavirus	Aparavirus kashmirens	Kashmir bee virus (KBV)	68876	162	-	97.5	116.3	Invertebrate
		Aphis gossypii virus	-	2448899	38	-	98.4	105	Invertebrate
		Cripavirus mortiferum	Aphid lethal paralysis virus (ALPV)	209529	5171	5	96.4	1730	Invertebrate
	Cripavirus ropadi	Rhopalosiphum padi virus (RHPV)	-	66834	3405	9	98.6	1428.4	Invertebrate
Genomoviridae	Gemykibivirus	Gemykibivirus humaS5	Human associated gemykibivirus 5 (HAGRIV-5)	2004962	537	2	93.4	530.1	Human
	Unclassified Genomoviridae	Tartoise genomovirus 3	-	2582886	28	1	84.3	115	Vertebrate
Ifloviridae	Iflovirus	Iflovirus aladeformis	Deformed wing virus (DWV)	198112	3466	16	98.7	417.9	Invertebrate
	Iflovirus	Iflovirus sacbroodi	Sacbrood virus (SBV)	89463	130	1	97.7	249.8	Invertebrate
		Iflovirus speixiguae	Spodoptera exigua iflavivirus 1 (SeIV-1)	1111710	19262	5	91.6	1101.6	Invertebrate
Nodaviridae	Alphanodavirus	Flock House virus	-	12287	230	2	87.5	381.9	Invertebrate
	Nodamura virus	-	-	12288	246	1	88.2	232.5	Invertebrate/Vertebrate
Parvoviridae	Bocaparvovirus	Canine bocavirus	-	1108949	66	1	91.6	418	Vertebrate
	Unclassified Parvoviridae	Murine-associated porcine bocavirus	-	2171393	34	-	97.3	150	Vertebrate
		Pigeon parvovirus A	-	1399967	324	3	94.8	186.6	Vertebrate
Picarnaviridae	Kobuvirus	Alchivirus A	Canine kobuvirus (CaCoV)	1836608	60	-	96.4	139.3	Vertebrate
	Megivirus	Megivirus B	Pigeon mesivirus 2 (Mev-B2)	1379161	580	3	94	295.6	Vertebrate
Powiridae	Leporipoxvirus	Myxoma virus	-	10273	30	-	100	132	Vertebrate
Virgaviridae	Tobamovirus	Rehmannia mosaic virus	-	425279	26	-	96.5	108	Plant
	Tobamovirus	Tobacco mosaic virus	-	12242	274	-	95.7	122.7	Plant
		Turnip vein-clearing virus	-	29272	32	-	94.6	139.3	Plant

Material suplementari de l'article 6

- **Table S1.** Quantification of HAdV in WWTP samples during winter and summer using qPCR. Concentrations in aerosol samples are expressed as GC/m³, while the surface values are expressed as GC/cm². Values have been remarked with colors for the samples used in sequencing pools. NA: Not analyzed, ND: Not Detected. NA and ND samples have not been included in sequencing pools.

(A continuació)

- **Table S2.** Quantification of PAdV in swine farm samples during winter and summer using qPCR. Concentrations in aerosol samples are expressed as GC/m³, while the surface values are expressed as GC/cm². Values have been remarked with colors for the samples used in sequencing pools. NA: Not analyzed, ND: Not Detected. NA and ND samples have not been included in sequencing pools.

(A continuació)

- **Table S3.** Vertebrate viruses sequenced in winter aerosol samples pooled from the WWTP centrifuge zone (indoor) using target enrichment sequencing.

(A continuació)

- **Table S4.** Vertebrate viruses sequenced in summer aerosol samples pooled from the WWTP centrifuge zone (indoor) using target enrichment sequencing.

(A continuació)

- **Table S5.** Vertebrate viruses sequenced in the summer aerosol sample from the WWTP reactor zone (outdoor) using target enrichment sequencing.

(A continuació)

- **Table S6.** Vertebrate viruses sequenced in winter surface samples pooled from the WWTP workspaces using target enrichment sequencing.

(A continuació)

- **Table S7.** Vertebrate viruses sequenced in winter surface samples pooled from the WWTP break room using target enrichment sequencing.

(A continuació)

- **Table S8.** Vertebrate viruses sequenced in summer surface samples pooled from the WWTP workspaces using target enrichment sequencing.

(A continuació)

- **Table S9.** Vertebrate viruses sequenced in summer surface samples pooled from the WWTP break room using target enrichment sequencing.

(A continuació)

- **Table S10.** Vertebrate viruses sequenced in winter aerosol samples pooled from the swine farm using target enrichment sequencing.

(A continuació)

- **Table S11.** Vertebrate viruses sequenced in summer aerosol samples pooled from the swine farm using target enrichment sequencing.

(A continuació)

- **Table S12.** Vertebrate viruses sequenced in winter surface samples pooled from the swine farm using target enrichment sequencing.

(A continuació)

- **Table S13.** Vertebrate viruses sequenced in summer surface samples pooled from the swine farm using target enrichment sequencing.

(A continuació)

Table S1. Quantification of HAdV in WWTP samples during winter and summer using qPCR. Concentrations in aerosol samples are expressed as GC/m3, while the surface values are expressed as GC/cm2. Values have been remarked with colors for the samples used in sequencing pools. NA: Not analyzed, ND: Not Detected. NA and ND samples have not been included in sequencing pools.

	Zone		Sample	Sample code	Winter season			Sequencing pools	Summer season			Sequencing pools		
					24/2/2022	10/3/2022	24/3/2022		8/6/2022	21/6/2022	6/7/2022		21/7/2022	
Aerosol (GC/m3)	Centrifuge (indoor)	Air E1	05A	5,18E+00	3,10E+00	6,34E+00	Pool 1	9,62E+01	8,61E+01	ND	ND	Pool 2		
	Reactor (outdoor)	Air E2	11AR		NA	NA	NA	9,04E-01	NA	ND	ND	Pool 3		
Surface (GC/cm2)	Workspaces	Filtering	Handle 1	01M1	5,34E+00	9,02E+00	4,74E-01	Pool 4	NA	6,78E+01	NA	NA	Pool 6	
			Handle 2	01M2	ND	3,08E+01	3,99E+01		NA	1,89E+01	ND	2,17E+01		
			Handrail	01B	NA	5,16E+00	5,71E+00		NA	NA	NA	NA		
			Handle 3	03M	1,20E+01	5,10E+02	3,11E+01		NA	3,85E+01	ND	6,50E+00		
		Centrifuge	Hose 1	03MP	1,05E+02	NA	NA		NA	NA	NA	NA		NA
			Hose 2	05MP	2,23E+02	1,54E+03	1,78E+02		NA	6,01E+01	ND	ND		
			Handle 4	05M	1,21E+02	3,55E+01	2,42E+01		NA	NA	NA	NA		
			Switch 1	09IL	ND	ND	ND		NA	NA	NA	NA		
		Main building - entrance	Switch 2	09IP	ND	1,19E+01	ND		NA	NA	NA	NA		
			Comptuer	09O	ND	2,41E+00	ND		NA	NA	NA	NA		
			Handle 6	09PE	7,37E-01	4,26E+00	ND		NA	ND	4,42E+00	1,72E+01		
	Reactor	Handle 7	11R	NA	6,60E+01	1,83E+00	NA	ND	ND	ND	NA			
		Break room	Coffee machine	07MC	ND	ND	ND	Pool 5	NA	NA	NA	NA	Pool 7	
			Vending machine	07MV	ND	8,45E+00	7,94E+00		NA	ND	3,06E+01	ND		
			Microwave	07M	ND	ND	3,37E+00		NA	NA	NA	NA		
			Toaster	07T	ND	ND	1,68E+00		NA	NA	NA	NA		
			Fridge	07N	6,08E+00	9,23E+01	7,07E+00		NA	3,81E+00	1,56E+01	9,58E-01		
			Handle 5	07PV	ND	ND	1,51E+01		NA	NA	NA	NA		

Table S2. Quantification of PAdV in swine farm samples during winter and summer using qPCR. Concentrations in aerosol samples are expressed as GC/m3, while the surface values are expressed as GC/cm2. Values have been remarked with colors for the samples used in sequencing pools. NA: Not Detected, ND: Not Detected. NA and ND samples have not been included in sequencing pools.

	Sample	Sample code	Winter season					Sequencing pools	Summer season			Sequencing pools
			4/2/2022	17/2/2022	3/3/2022	22/3/2022	22/3/2022		8/7/2022	18/7/2022	28/7/2022	
Aerosol (GC/m3)	Air F	0AF	1,53E+00	2,08E+01	1,27E+02	NA	NA	Pool 8	ND	6,15E+00	1,28E+02	Pool 9
Surface (GC/cm2)	Front door handle	02P	NA	4,39E+01	1,10E+01	5,45E+01	5,45E+01	Pool 10	9,30E+04	6,34E+04	5,52E+02	Pool 11
	Switch 1	02I1	NA	ND	ND	2,41E+01	2,41E+01		3,12E+04	1,05E+04	1,65E+02	
	Switch 2	02I2	NA	5,50E+01	ND	3,81E+02	3,81E+02		1,10E+04	5,84E+03	3,43E+02	
	Breeding door handle	04P	NA	ND	ND	8,06E+01	8,06E+01		1,94E+05	1,14E+05	8,88E+02	
	Lateral door handle	04PD	NA	3,12E+02	1,12E+02	3,34E+02	3,34E+02		7,19E+04	2,96E+04	8,16E+01	
	Wheelbarrow	04C	NA	NA	7,61E+01	NA	NA		NA	NA	NA	
	Switch 3	06I	NA	1,92E+02	2,54E+02	7,13E+02	7,13E+02		1,88E+05	2,17E+05	9,00E+02	

Table S3. Vertebrate viruses sequenced in winter aerosol samples pooled from the WWTP centrifuge zone (indoor) using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average lenght (bp)	Host
Adenoviridae	Mastadenovirus	Human mastadenovirus F	HAdV-41 (Human adenovirus 41)	130309	66	-	100	133.6	Human
	Gyrovirus	Gyrovirus chickenanemia	CAV (Chicken anemia virus)	12618	24	-	99	147	Bird
Parvoviridae	Bocaparvovirus	Primate bocaparvovirus 1	HBoV3 (Human bocavirus 3)	1511871	66	1	99.6	235.7	Human
		Primate bocaparvovirus 2	-	1511872	26	-	99.6	266	Human
	Dependoparvovirus	Adeno-associated virus	-	272636	1064	5	98.3	549.8	Vertebrate
		Avian dependoparvovirus 1	-	1511894	20	-	98.7	128	Bird
Picornaviridae	Siccinivirus	Siccinivirus A	-	1755588	32	1	87.5	236.8	Bird

Table S4. Vertebrate viruses sequenced in summer aerosol samples pooled from the WWTP centrifuge zone (indoor) using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average length (bp)	Host
Adenoviridae	Aviadenovirus	<i>Fowl aviadenovirus D</i>	-	190064	322	2	99.8	249.3	Bird
		<i>Fowl aviadenovirus E</i>	-	190065	106	1	100	274.3	Bird
	Mastadenovirus	<i>Human mastadenovirus B</i> <i>Human mastadenovirus F</i>	- HAdV-41 (Human adenovirus 41)	108098 130309	28 364	- 1	100 99.6	141 199.5	Human Human
Anelloviridae	Gyrovirus	<i>Gyrovirus chickenanemia</i> <i>Gyrovirus galga1</i>	CAV (Chicken anemia virus) AGV2 (Avian gyrovirus 2)	12618 1002273	884 280	2 1	99.2 98.1	1176.4 273.4	Bird Bird
		<i>Gyrovirus hmsa3</i>	GyV4 (Gyrovirus 4)	1214955	114	1	96.6	327.5	Human/vertebrate
	Mastastrovirus	<i>Human astrovirus MLB1</i>	-	568715	33	-	99.6	135.8	Human
Astroviridae	Unclassified Astroviridae	<i>Astrovirus VA3</i>	-	683174	40	-	98.8	149	Human
		<i>Human astrovirus</i>	-	1868658	54	-	98.7	125.4	Human
Caliciviridae	Norovirus	<i>Norwalk virus</i>	-	11983	26	-	99.3	142.5	Human/vertebrate
Circoviridae	Circovirus	<i>Circovirus VS6600032</i>	-	1525174	76	-	99.1	127.8	Human
		<i>Human circovirus VS6600022</i>	-	1525173	36	1	98.6	285	Human
		<i>Pigeon circovirus</i>	-	1414603	48	-	97.8	123.4	Bird
	Cyclovirus	<i>Sewage associated circovirus</i>	-	689405	56	-	95	137.7	-
		<i>Cyclovirus bashri</i> <i>Cyclovirus TN10</i>	HuACyV4 (Human associated cyclovirus 4) -	2038722 742955	34 111	- 1	97.8 98.9	137 350	Human Human
Coronaviridae	Gammacoronavirus	<i>Avian coronavirus</i>	IBV (Avian infectious bronchitis virus)	694014	534	5	99.1	243.7	Bird
Orthoherpesviridae	Mardivirus	<i>Mardivirus gallidalpha2</i>	GaHV-2 (Gallid alphaherpesvirus 2)	10390	88	-	99.7	143.1	Bird
		<i>Carnivore bocaparvovirus 1</i>	CPV-1 (Canine parvovirus 1)	1511866	30	-	100	114	Dog
Parvoviridae	Bocaparvovirus	<i>Human bocavirus</i>	HBoV-CUK 18 (Human bocavirus strain CUK18)	329641	74	1	99.1	221	Human
		<i>Porcine bocavirus</i>	-	1165907	28	-	96.7	151	Pig
		<i>Primate bocaparvovirus 1</i>	HBoV3 (Human bocavirus 3)	1511871	249	2	99.6	312.4	Human
	Dependoparvovirus	<i>Primate bocaparvovirus 2</i>	-	1511872	123	-	99.3	154.7	Human
		<i>Adeno-associated virus</i>	-	272636	4036	7	98.7	1179.6	Vertebrate
		<i>Avian dependoparvovirus 1</i>	-	1511894	2038	9	96.8	524	Bird
	Protoparvovirus	<i>Carnivore protoparvovirus 1</i>	-	1511906	20	-	99.8	141	Vertebrate
		<i>Rat minute virus 1</i>	-	344593	278	2	98.2	449.6	Rat
		<i>Rat parvovirus NTU1</i>	-	1248250	73	1	98.6	206.1	Rat
		<i>Rodent protoparvovirus 2</i>	-	1511909	15	-	98.3	140.7	Rodent
Picornaviridae		<i>Ungulate protoparvovirus 1</i>	PPV1 (Porcine parvovirus 1)	1511910	58	1	99.7	206.8	Pig
	Cardiovirus	<i>Cardiovirus B</i>	-	1821750	40	-	94.2	144.3	Human/vertebrate
	Kobuvirus	<i>Norway rat kobuvirus 1</i>	-	1562048	92	1	96.6	292	Rat
	Megrovirus	<i>Megrovirus C</i>	ChMV (Chicken megrovirus)	1477515	283	2	93.3	224.8	Bird
	Unclassified Picornaviridae	<i>Chicken proventriculitis virus</i>	-	1566005	35	-	93.4	137	Bird

Table S5. Vertebrate viruses sequenced in the summer aerosol sample from the WWTP reactor zone (outdoor) using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average lenght (bp)	Host
Adenoviridae	Aviadenovirus	Fowl aviadenovirus D	-	190064	18	-	97.8	135	Bird
	Mastadenovirus	Human mastadenovirus F	HAdV-41 (Human adenovirus 41)	130309	206	2	97.7	219.8	Human
	Mastrovirus	Human astrovirus MLB1	-	568715	33	-	99.6	135.8	Human
Caliciviridae	Norovirus	Norwalk virus	-	11983	68	-	99.7	140.2	Human/vertebrate
Circoviridae	Circovirus	Circovirus porcine2	PCV-2 (Porcine circovirus 2)	85708	3070	1	99.9	783	Pig
	Betapapillomavirus	Betapapillomavirus 3	-	334207	22	-	98.7	115	Human
Parvoviridae	Bocaparvovirus	Carnivore bocaparvovirus 1	CPV-1 (Canine parvovirus 1)	1511866	42	1	99.6	244	Dog
	Dependoparvovirus	Adeno-associated virus	-	272636	34	-	99.6	124	Vertebrate
	Protaparvovirus	Rat minute virus 1	-	344593	236	1	99	494	Rat

Table S6. Vertebrate viruses sequenced in winter surface samples pooled from the WWTP workspaces using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average length (bp)	Host
Adenoviridae	Avidenovirus	Fowl aviadenovirus D	-	190064	236	1	99.8	247.2	Bird
		Fowl aviadenovirus E	-	190065	492	1	99.7	715.3	Bird
	Mastadenovirus	Human mastadenovirus A	-	129875	26	-	98.7	152	Human
		Human mastadenovirus C	-	129951	356	2	100	494.7	Human
		Human mastadenovirus F	HAdV-41 (Human adenovirus 41)	130809	12359	17	99.8	902.7	Human
Circoviridae	Circovirus	Circovirus VS6600032	-	1525174	366	2	97.7	403.1	Human
Orthoherpesviridae	Lymphocryptovirus	Lymphocryptovirus humangamma4	EBV (Epstein-Barr virus)	10376	36	1	100	299	Human
	Mardivirus	Mardivirus gallidalpha2	GaHV-2 (Gallid alphaherpesvirus 2)	10390	2272	2	99.1	373.5	Bird
		Mardivirus gallidalpha3	GaHV-3 (Gallid alphaherpesvirus 3)	35250	424	1	99.9	368.9	Bird
Papillomaviridae	Alphapapillomavirus	Alphapapillomavirus 4	-	337043	26	-	99.1	150	Human
	Betapapillomavirus	Betapapillomavirus 3	-	334207	56	-	99.4	114.7	Human
	Chipapillomavirus	Chipapillomavirus 2	CPV-4 (Canine papillomavirus 4)	1177759	1162	4	99.9	487.2	Dog
	Gamma papillomavirus	Gamma papillomavirus 10	HPV-133 (Human papillomavirus 133)	1175852	570	1	96.5	486.6	Human
		Gamma papillomavirus 21	HPV-167 (Human papillomavirus 167)	1961678	96	1	98.8	325	Human
	Aveparvovirus	Galliform aveparvovirus 1	-	1511865	249	1	97	290	Bird
Parvoviridae	Bocaparvovirus	Carnivore bocaparvovirus 1	CPV-1 (Canine parvovirus 1)	1511866	190	1	99	400.6	Dog
		Murine-associated porcine bocavirus	-	2171393	22	-	97.8	106	Pig
		Primate bocaparvovirus 1	HBov3 (Human bocavirus 3)	1511871	1858	1	98.8	637.4	Human
	Dependoparvovirus	Rodent bocavirus	-	2137546	44	-	98.9	128.2	Rodent
		Adeno-associated virus	-	272636	44816	11	98.4	874.3	Vertebrate
Polyomaviridae	Alphapolyomavirus	Avian dependoparvovirus 1	-	1511894	661	2	95.8	632.8	Bird
Retroviridae	Alpharetrovirus	Human polyomavirus 5	MCPyV (Merkel cell polyomavirus)	1891726	292	2	99.4	237.6	Human
	Alpharetrovirus	Avian leukosis virus	-	11864	15399	1	99.7	615.9	Bird

Table S7. Vertebrate viruses sequenced in winter surface samples pooled from the WWTP break room using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average length (bp)	Host
Adenoviridae	Aviadenovirus	Turkey aviadenovirus D	(TAdV-4) Turkey adenovirus 5	1962302	6232	1	99.4	500.2	Bird
Herpesviridae	Macavirus	Macavirus suldtamma5	PLHV3 (Porcine lymphotropic herpesvirus 3)	1960251	70	1	99.7	349	Pig
Orthoherpesviridae	Lymphocryptovirus	Lymphocryptovirus humangamma4	EBV (Epstein-Barr virus)	10376	3813	4	99.9	802	Human
	Roseolovirus	Human betaherpesvirus 7	HHV-7 RK (Human herpesvirus 7 strain RK)	10372	574	2	99.6	287.5	Human
Papillomaviridae	Betapapillomavirus	Betapapillomavirus 1	HPV-20 (Human papillomavirus 20)	337051	296	1	99.6	213.4	Human
		Betapapillomavirus 2	HPV-80 (Human papillomavirus 80)	333924	36610	4	99.6	1461.3	Human
	Gammapapillomavirus	Gammapapillomavirus 7	HPV-123 (Human papillomavirus 123)	1175849	2412	2	98.6	865.2	Human
		Gammapapillomavirus 18	HPV-156 (Human papillomavirus 156)	1513263	4388	1	99.3	762	Human
		Gammapapillomavirus 22	HPV-172 (Human papillomavirus 172)	1961679	3478	1	98.5	589	Human
		Gammapapillomavirus 23	HPV-175 (Human papillomavirus 175)	1961680	798	2	98.5	434.8	Human
Mupapillomavirus	Mupapillomavirus 3	-	1961783	28	-	99.7	142.5	Human	
Retroviridae	Alpharetrovirus	Avian leukosis virus	-	11864	9138	6	98.3	437.6	Bird
	Unclassified Retroviridae	Avian endogenous retrovirus EAV-HP	-	93465	4929	4	99.1	374.1	Bird

Table S8. Vertebrate viruses sequenced in summer surface samples pooled from the WWTP workspaces using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average length (bp)	Host
Adenoviridae	Aviadenovirus	Fowl aviadenovirus D	-	190064	356	1	99.7	338.3	Bird
	Mastadenovirus	Human mastadenovirus F	HAAdV-41 (Human adenovirus 41)	130309	356	1	98.8	303.3	Human
Papillomaviridae	Betapapillomavirus	Betapapillomavirus 1	HPV-98 (Human papillomavirus 98)	337051	1224	7	98.8	398.6	Human
		Betapapillomavirus 2	HPV-15 (Human papillomavirus 15)	333924	222	1	99.5	143	Human
		Betapapillomavirus 3	HPV-115 (Human papillomavirus 115)	334207	2323	5	98.8	473.2	Human
	Chipapillomavirus	Chipapillomavirus 2	CPV-4 (Canine papillomavirus 4)	1177759	74	1	97.3	208.7	Dog
		Gammapapillomavirus 1	HPV-65 (Human papillomavirus 65)	333926	2942	6	98.9	519.3	Human
	Gammapapillomavirus	Gammapapillomavirus 7	-	1175849	118	1	99.2	220.2	Human
		Gammapapillomavirus 15	HPV-135 (Human papillomavirus 135)	1513260	98	1	88.4	247.7	Human
		Gammapapillomavirus 21	HPV-167 (Human papillomavirus 167)	1961678	864	1	99.1	665	Human
Parvoviridae	Bocaparvovirus	Ungulate bocaparvovirus 5	PBoV3 (Porcine bocavirus 3)	1511877	20	-	95.3	127	Pig

Table S9. Vertebrate viruses sequenced in summer surface samples pooled from the WWTP break room using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average length (bp)	Host
Circoviridae	Cyclovirus	Cyclovirus bashiri	HuACYV4 (Human associated cyclovirus 4)	2038722	400	2	96.9	372.4	Human
	Lymphocryptovirus	Lymphocryptovirus humangamma4	EBV (Epstein-Barr virus)	10376	8849	4	99.6	808.8	Human
	Roseolavirus	Human betaherpesvirus 6	HHV-6 NY-40 (Human herpesvirus 6 strain NY-40)	10368	371	1	99.6	532.9	Human
		Human betaherpesvirus 7	HHV-7 RK (Human herpesvirus 7 strain RK)	10372	8659	10	99.6	1040.8	Human
	Alphapapillomavirus	Alphapapillomavirus 2	HPV-3 (Human papillomavirus 3)	337039	2869	2	99.4	557	Human
		Betapapillomavirus 1	HPV-36 (Human papillomavirus 36)	337051	44485	29	99.1	875.3	Human
		Betapapillomavirus 2	HPV-80 (Human papillomavirus 80)	333924	16217	14	98.7	811.3	Human
		Betapapillomavirus 3	HPV-75 (Human papillomavirus 75)	334207	28022	14	98.4	2263.4	Human
		Betapapillomavirus 5	HPV-96 (Human papillomavirus 96)	334209	3902	7	98.5	646.3	Human
	Papillomaviridae	Chipapillomavirus 2	CPV-4 (Canine papillomavirus 4)	1177759	188	1	99.5	146	Dog
		Gammapapillomavirus 3	-	333929	16	-	100	119	Human
		Gammapapillomavirus 7	-	1175849	2638	6	96.7	533	Human
		Gammapapillomavirus 8	HPV-164 (Human papillomavirus 164)	1175850	6696	4	97	646.8	Human
		Gammapapillomavirus 10	-	1175852	122	-	99.5	136.6	Human
Retroviridae	Gammapapillomavirus	Gammapapillomavirus 12	HPV-195 (Human papillomavirus 199)	1513257	1038	2	99.6	569	Human
		Gammapapillomavirus 13	HPV-128 (Human papillomavirus 128)	1513258	448	2	99	265.5	Human
		Gammapapillomavirus 15	HPV-135 (Human papillomavirus 135)	1513260	126	1	99.4	230.9	Human
		Alpharetrovirus	-	11864	642	2	94.8	235.8	Bird
		Betaretrovirus	-	2762664	226	1	96.7	484	Sheep and goats
	Betaretrovirus	Endogenous enzootic nasal tumor virus	-	2584748	78	-	99.8	112	Sheep and goats
		Enzootic nasal tumour virus 2	-						

Table S10. Vertebrate viruses sequenced in winter aerosol samples pooled from the swine farm using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average length (bp)	Host
Adenoviridae	Mastadenovirus	Porcine adenovirus C	PAdV-5 (Porcine adenovirus 5)	130371	252	2	99.2	182 Pig	
Anelloviridae	Isotatorquevirus	Torque teno sus virus 1b	-	687387	176	1	99.5	273.4 Pig	
Astroviridae	Mamastrovirus	Feline astrovirus (Mamastrovirus 2)	-	1239566	930	1	90.8	953.5 Vertebrates	
		Mamastrovirus 3	-	1239567	2534	9	92.6	599 Vertebrates	
		Mamastrovirus pig/HL/China	-	1562563	52	1	89.8	275 Pig	
		Porcine astrovirus 2	-	1105378	367	4	93.7	212.6 Pig	
		Porcine astrovirus 4	-	1105379	510	1	95.1	473.2 Pig	
Unclassified Astroviridae	Porcine astrovirus 5	Porcine astrovirus 5	-	1093928	100	1	96.5	309.2 Pig	
		Circovirus porcine2	PCV-2 (Porcine circovirus 2)	85708	104	1	99.1	319 Pig	
Circoviridae	Unclassified Circoviridae	Fur seal faeces associated circular DNA virus	-	1353241	14518	4	92	1237.2 Fur seals and pigs	
		Po-Circo-like virus	-	1454552	1730	5	92.7	499.8 Pig	
		Po-Circo-like virus 41	-	1105385	136	1	95.1	318.3 Pig	
		Po-Circo-like virus S1	-	1105386	101	2	95	175.3 Pig	
		Porcine stool-associated circular virus 5	-	1475063	5901	11	95.6	342.7 Pig	
Genomoviridae	Gemykibivirus	Porcine stool-associated circular virus/BEL/15V010	-	2017724	438	1	93.7	418.7 Pig	
		Gemykibivirus anima1	PfAgV (Porcine feces-associated gemykibircularvirus)	2017712	16	-	95.8	261 Pig	
		Macavirus suisdamm05	PLHV3 (Porcine lymphotropic herpesvirus 3)	1960251	50	1	100	326 Pig	
Herpesviridae	Macavirus	Macavirus suisdamm05	PLHV3 (Porcine lymphotropic herpesvirus 3)	1177764	2061	8	99.9	526.9 Pig	
Papillomaviridae	Dyodeltapapillomavirus	Dyodeltapapillomavirus 1	SsPV1 (Sus scrofa papillomavirus 1)	1177764	2061	8	99.9	526.9 Pig	
		Porcine bocavirus	-	1165907	12180	18	93.8	745.5 Pig	
		Porcine bocavirus 4	-	1084716	932	5	94.6	337.2 Pig	
		Porcine bocavirus 5	-	1133653	798	3	94.1	376.9 Pig	
		Rhinolophus sinicus bat bocapapovirus 2	-	2053081	219	1	94.1	244.3 Bat	
Parvoviridae	Bocapapovirus	Ungulate bocapapovirus 2	PBoV1 (Porcine bocavirus 1)	1511874	2390	3	94	725 Pig	
		Ungulate bocapapovirus 5	PBoV3 (Porcine bocavirus 3)	1511877	52160	43	93.2	1452.5 Pig	
		Ungulate capipapovirus 2	PPV4 (Porcine parvovirus 4)	1511890	72	1	99.6	225.4 Pig	
		Adeno-associated virus	-	272636	56	-	93.4	135.8 Vertebrate	
		Adeno-associated virus - Po1	-	607682	100	1	97.4	153 Pig	
Picornaviridae	Kobuvirus	Adeno-associated virus - Po3	-	607684	18	-	99	102 Pig	
		Aichivirus C	PkoV (Porcine kobuvirus)	1298633	308	2	92.9	415.2 Pig	
Retroviridae	Sagelovirus	Porcine enterovirus A	-	686984	96	-	94.9	112.4 Pig	
		Porcine type-C oncovirus	-	369960	16256	17	98.6	1190.3 Pig	
Tobamviridae	Torovirus	Bovine torovirus	-	74501	2402	4	94.3	395.3 Cattle	
		Porcine torovirus	-	237020	13874	21	94.7	909.9 Pig	

Table S11. Vertebrate viruses sequenced in summer aerosol samples pooled from the swine farm using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average length (bp)	Host
Anelloviridae	Kappatorquevirus	<i>Kappatorquevirus suidak2a</i>	TTSuVk2a (Torque teno sus virus k2a)	1968861	28	1	96.8	249 Pig	
		<i>Feline astrovirus (Mamastrovirus 2)</i>	-	1239566	538	1	93.1	712.9 Vertebrates	
Astroviridae	Mamastrovirus	<i>Mamastrovirus 3</i>	-	1239567	444	3	96.3	222.3 Vertebrates	
		<i>Mamastrovirus pig/HL/China</i>	-	1562563	20	-	94.2	113 Pig	
		<i>Porcine astrovirus 2</i>	-	1105378	145	1	92.7	157.3 Pig	
		<i>Porcine astrovirus 4</i>	-	1105379	67	-	97.1	104.6 Pig	
Circoviridae	Unclassified Circoviridae	<i>Fur seal faeces associated circular DNA virus</i>	-	1353241	496	4	91.7	394.3 Fur seals and pigs	
		<i>Po-Circo-like virus</i>	-	1454552	1724	6	90.3	457.3 Pig	
		<i>Po-Circo-like virus 41</i>	-	1105385	5463	9	96.8	194,75 Pig	
		<i>Po-Circo-like virus 51</i>	-	1105386	1163	2	93.7	714.3 Pig	
		<i>Porcine stool-associated circular virus 5</i>	-	1475063	122	1	96.7	193.2 Pig	
		<i>Porcine stool-associated circular virus/BEL/15V010</i>	-	2017724	607	1	92	274.6 Pig	
		<i>Macavirus suidgamma5</i>	PLHV3 (Porcine lymphotropic herpesvirus 3)	1960251	1124	6	99.6	184.8 Pig	
Herpesviridae	Macavirus								
Papillomaviridae	Dyodeltapapillomavirus	<i>Dyodeltapapillomavirus 1</i>	SsPV1 (Sus scrofa papillomavirus 1)	1177764	248	2	99.6	187.3 Pig	
		<i>Porcine bocavirus</i>	-	1165907	4080	14	93.1	1209.9 Pig	
Parvoviridae	Bocaparvovirus	<i>Porcine bocavirus 4</i>	-	1084716	249	3	93.5	383 Pig	
		<i>Porcine bocavirus 5</i>	-	1133653	19	-	91.9	185 Pig	
		<i>Rhinolophus sinicus bat bocaparvovirus 2</i>	-	2053081	69	1	89.5	188.8 Bat	
		<i>Ungulate bocaparvovirus 2</i>	PBoV1 (porcine bocavirus 1)	1511874	183	-	97	142.6 Pig	
		<i>Ungulate bocaparvovirus 3</i>	PBoV5 (Porcine bocavirus 5)	1511875	17	-	98.4	127 Pig	
		<i>Ungulate bocaparvovirus 5</i>	PBoV3 (porcine bocavirus 3)	1511877	9145	22	93.6	872.5 Pig	
		<i>Aichivirus C</i>	PkoV (porcine kobuvirus)	1298633	248	2	92.9	343.4 Pig	
Picornaviridae	Kobuvirus								
Retroviridae	Gammaretrovirus	<i>Porcine type-C oncovirus</i>	-	369960	8770	21	98.7	774.1 Pig	
		<i>Bovine torovirus</i>	-	74501	358	3	94.6	270.3 Cattle	
Tobamiviridae	Torovirus	<i>Porcine torovirus</i>	-	237020	1334	9	95.8	282.9 Pig	

Table S12. Vertebrate viruses sequenced in winter surface samples pooled from the swine farm using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average length (bp)	Host
Adenoviridae	Mastadenovirus	Porcine adenovirus C	PAqV-5 (Porcine adenovirus 5)	130371	4953	21	98.9	600.4	Pig
Circoviridae	Circovirus	Circovirus porcine1	PCV-1 (Porcine circovirus 1)	133704	2530	1	99.3	841.3	Pig
		Circovirus porcine3	PCV-3 (Porcine circovirus 3)	1868221	5576	1	99.7	1139	Pig
		Fur seal faeces associated circular DNA virus	-	1353241	3886	4	93.1	1166.9	Fur seals and pigs
	Unclassified Circoviridae	Po-Circo-like virus 41	-	1105385	272	1	95.3	613.5	Pig
		Po-Circo-like virus 51	-	1105386	510	2	90.5	584.1	Pig
Herpesviridae	Macavirus	Macavirus suisgamma5	PLHV3 (porcine lymphotropic herpesvirus 3)	1960251	6510	18	99.5	619.7	Pig
	Roseolavirus	Roseolavirus suisbeta2	SHV-2 (Porcine cytomegalovirus)	1608255	12807	30	99.4	323.3	Pig
Papillomaviridae	Dyodeltapapillomavirus	Dyodeltapapillomavirus 1	SsPV1 (Sus scrofa papillomavirus 1)	1177764	161	1	99.7	298.2	Pig
Parvoviridae	Bocaparvovirus	Porcine bocavirus	-	1165907	19983	5	98.8	3968.2	Pig
		Porcine bocavirus 4	-	1084716	169	1	97.5	756	Pig
		Porcine bocavirus 5	-	1133653	83	1	93.3	239	Pig
		Ungulate bocaparvovirus 2	PBoV1 (Porcine bocavirus 1)	1511874	218	1	97.5	199	Pig
		Ungulate bocaparvovirus 5	PBoV3 (Porcine bocavirus 3)	1511877	5327	16	91.6	532.3	Pig
Picornaviridae	Kabuvirus	Aichivirus C	PKoV (Porcine kobuvirus)	1298633	168	2	94.4	297.9	Pig
Polyomaviridae	Unclassified Polyomaviridae	Porcine polyomavirus	-	2126099	116	1	99.6	355	Pig
Retroviridae	Gammaretrovirus	Porcine type-C oncovirus	-	369960	1362922	466	97.9	905.9	Pig

Table S13. Vertebrate viruses sequenced in summer surface samples pooled from the swine farm using target enrichment sequencing.

Family	Genus	Species	Main Genogroup/Genotype/Serotype	TaxID	Reads	Contigs	Average Nucleotide Identity (%)	Average length (bp)	Host
Adenoviridae	Mastadenovirus	Porcine mastadenovirus A	PAdV-3 (Porcine adenovirus 3)	130369	36	-	93.6	105.8 Pig	
		Porcine mastadenovirus C	PAdV-5 (Porcine adenovirus 5)	130371	834	1	99.05	343.9 Pig	
Astroviridae	Mamastrovirus	Mamastrovirus 3	-	1239567	509	4	93.7	370.7 Vertebrates	
		Porcine astrovirus 2	-	1105378	169	1	90.9	209.7 Pig	
		Porcine astrovirus 4	-	1105379	89	-	92.1	132.6 Pig	
Circoviridae	Unclassified Circoviridae	Circovirus sp.	-	1964372	1745	3	93.2	530.9 Vertebrates	
		Fur seal faeces associated circular DNA virus	-	1353241	9042	4	90.5	1273 Fur seals and pigs	
		Po-Circo-like virus	-	1454552	1328	2	92.5	833.5 Pig	
		Po-Circo-like virus 41	-	1105385	625	2	96.2	769.9 Pig	
		Po-Circo-like virus 51	-	1105386	547	2	93.5	505.5 Pig	
		Porcine stool-associated circular virus 5	-	1475063	320	2	91.1	301.8 Pig	
		Porcine stool-associated circular virus/BEL/15V010	-	2017724	58	-	96.2	111 Pig	
		Macavirus suis/gamma5	PLHV3 (Porcine lymphotropic herpesvirus 3)	1960251	74	- 99.5	-	120.6 Pig	
		Roseolavirus suis/delta2	SHV-2 (Porcine cytomegalovirus)	1608255	12807	30	99.4	323.3 Pig	
		Dyodeltapapillomavirus 1	SSPV1 (Sus scrofa papillomavirus 1)	1177764	276	1	99.6	1014 Pig	
Papillomaviridae	Dyodeltapapillomavirus	Gammagapillomavirus 24	HPV178 (Human papillomavirus 178)	1961681	1466	1	97.8	1058 Human	
Parvoviridae	Bocaparvovirus	Porcine bocavirus	-	1165907	21350	9	96.9	1100.5 Pig	
		Rhinolophus shicus bat bocaparvovirus 2	-	2053081	1519	1	94.8	822 Bat	
		Ungulate bocaparvovirus 2	PBoV1 (Porcine bocavirus 1)	1511874	258	1	98.7	285 Pig	
		Ungulate bocaparvovirus 5	PBoV3 (Porcine bocavirus 3)	1511877	48161	9	93.8	1100.7 Pig	
		Ungulate capiparvovirus 2	PPV4 (Porcine parvovirus 4)	1511890	26	-	99.3	148.5 Pig	
		Porcine parvovirus 5	-	1241957	146	1	100	379 Pig	
		Ungulate tetraparvovirus 3	PPV2 (Porcine parvovirus 2)	1126383	138	1	98.3	294 Pig	
Picornaviridae	Kobuvirus	Enterovirus G	EV-G1 (Porcine enterovirus G1)	106966	58	1	92.4	210 Pig	
		Achivirus C	PKoV (Porcine kobuvirus)	1298633	64528	14	91.6	1538.3 Pig	
Polyomaviridae	Unclassified Polyomaviridae	Porcine polyomavirus	-	2126099	809	3	99.6	307.4 Pig	
Retroviridae	Gammaretrovirus	Porcine type-C oncavirus	-	369560	602844	183	98.4	786.9 Pig	
Sedoreoviridae	Rotavirus	Rotavirus A	(PorV) Porcine rotavirus A	28875	822	1	94.3	405 Pig	

Altres publicacions

A continuació es detallen, per ordre cronològic, altres publicacions no incloses en aquesta tesi però realitzades durant el doctorat:

- Rusiñol, M., Martínez-Puchol, S., Forés, E., **Itarte, M.**, Girones, R. & Bofill-Mas, S. (2020). Concentration methods for the quantification of coronavirus and other potentially pandemic enveloped virus from wastewater. *Current Opinion in Environmental Science & Health*. 17, 21-28.
<https://doi.org/10.1016/j.coesh.2020.08.002>

Factor d'impacte: 8.1; Q1, D2 (Environmental Sciences).

- Forés, E., Bofill-Mas, S., **Itarte, M.**, Martínez-Puchol, S., Hundesa, A., Calvo, M., Borrego, C. M., Corominas, LI., Girones, R. & Rusiñol, M. (2021). Evaluation of two rapid ultrafiltration-based methods for SARS-CoV-2 concentration from wastewater. *Science of The Total Environment*. 768, 144786.
<https://doi.org/10.1016/j.scitotenv.2020.144786>

Factor d'impacte: 10.754; Q1, D1 (Environmental Sciences).

- Rusiñol, M., Zammit, I., **Itarte, M.**, Forés, E., Martínez-Puchol, S., Girones, R., Borrego, C., Corominas, LI. & Bofill-Mas, S. (2021). Monitoring waves of the COVID-19 pandemic: Inferences from WWTPs of different sizes. *Science of The Total Environment*, 787, 147463.
<https://doi.org/10.1016/j.scitotenv.2021.147463>

Factor d'impacte: 10.754; Q1, D1 (Environmental Sciences).

- Forés, E., Rusiñol, M., **Itarte, M.**, Martínez-Puchol, S., Calvo, M. & Bofill-Mas, S. (2022). Evaluation of a virus concentration method based on ultrafiltration and wet foam elution for studying viruses from large-volume water samples. *Science of The Total Environment*, 829, 154431.
<https://doi.org/10.1016/j.scitotenv.2022.154431>

Factor d'impacte: 9.8; Q1, D1 (Environmental Sciences).

- Guerrero-Latorre, L., Collado, N., Abasolo, N., et al. (2022). The Catalan Surveillance Network of SARS-CoV-2 in Sewage: design, implementation, and performance. *Scientific Reports*, 12, 16704. <https://doi.org/10.1038/s41598-022-20957-3>

Factor d'impacte: 4.6; Q2, D4 (Multidisciplinary sciences).

- Forés, E., Mejías-Molina, C., Ramos, A., **Itarte, M.**, Hundesa, A., Rusiñol, M., Martínez-Puchol, S., Esteve-Bricullé, P., Espejo-Valverde, A., Sirés, I., Calvo, M., Araujo, R. M. & Girones, R. (2023). Evaluation of pathogen disinfection efficiency of electrochemical advanced oxidation to become a sustainable technology for water reuse. *Chemosphere*. 313, 137393. <https://doi.org/10.1016/j.chemosphere.2022.137393>

Factor d'impacte: 8.8; Q1, D2 (Environmental Sciences).

- Mejías-Molina, C., Pico-Tomàs, A., Martínez-Puchol, S., **Itarte, M.**, Torrell, H., Canela, N., Borrego, C. M., Corominas, Ll., Rusiñol, M. & Bofill-Mas, S. (2024). Wastewater-based epidemiology applied at the building-level reveals distinct virome profiles based on the age of the contributing individuals. *Human Genomics*, 18, 10. <https://doi.org/10.1186/s40246-024-00580-1>

Factor d'impacte: 4.5; Q1, D3 (Genetics & Heredity).

Concentration methods for the quantification of coronavirus and other potentially pandemic enveloped virus from wastewater

Marta Rusiñol^{1,2}, Sandra Martínez-Puchol^{1,3}, Eva Forés^{1,3},
Marta Itarte^{1,3}, Rosina Girones^{1,3} and Sílvia Bofill-Mas^{1,3}

Abstract

As the novel SARS-CoV-2 was detected in faeces, environmental researchers have been using centrifugal ultrafiltration, polyethylene glycol precipitation and aluminium hydroxide flocculation to describe its presence in wastewater samples. High recoveries (up to 65%) are described with electronegative filtration when using surrogate viruses, but few literature reports recovery efficiencies using accurate quantification of enveloped viruses. Considering that every single virus will have a different behaviour during viral concentration, it is recommended to use an enveloped virus, and if possible, a betacoronavirus as murine hepatitis virus, as a surrogate. In this review, we show new data from a newly available technology that provides a quick ultrafiltration protocol for SARS-CoV-2. Wastewater surveillance is an efficient system for the evaluation of the relative prevalence of SARS-CoV-2 infections in a community, and there is the need of using reliable concentration methods for an accurate and sensitive quantification of the virus in water.

Addresses

¹ Section of Microbiology, Virology and Biotechnology, Department of Genetics, Microbiology and Statistics. Faculty of Biology, University of Barcelona, Catalonia, Spain

² GHS, Institute of Environmental Assessment & Water Research (IDAEA), CSIC, Barcelona, Spain

³ The Water Institute of the University of Barcelona, Spain

Corresponding author: Rusiñol, Marta (martarusinol@gmail.com)

Current Opinion in Environmental Science & Health 2020, 17:21–28

This review comes from a themed issue on **Environmental Health: COVID-19**

Edited by **Avelino Núñez-Delgado**

For a complete overview see the [Issue](#) and the [Editorial](#)

<https://doi.org/10.1016/j.coesh.2020.08.002>

2468-5844/© 2020 Elsevier B.V. All rights reserved.

Keywords

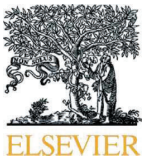
Concentration methods, Enveloped virus, SARS-CoV-2, Recovery efficiency, Surrogate virus.

Introduction

Many viruses that infect humans are excreted in large amounts through faeces, urine or skin desquamation, contributing to wastewater virome. Wastewater is a complex matrix that comprises a large variety of pathogenic and commensal viruses and provides important information about virus circulation, the introduction of emergent viruses and how they are transmitted among the population [1]. Waterborne viruses are generally nonenveloped and excreted in high numbers by infected individuals with or without disease, and in some cases long after the resolution of symptoms [2]. The study of excreted viruses is a very useful tool known as wastewater-based epidemiology, which has the potential to act as a complementary approach for current infectious disease surveillance systems and an early warning system for disease outbreaks [3].

The incidence of emerging microbes is a serious health concern worldwide. The increase of human–livestock contacts [4], population mobility and trade networks [5,6], climate change [7] or the wild meat trade and loss of animal habitats [8] has raised the risk of a global pandemics. Since 1980, nearly 90 novel human pathogen species have been discovered, more than 70 of those corresponded to novel human viruses, that compared to other pathogens have the potential to evolve more rapidly, being 80 of these associated with nonhuman reservoirs [9,10]. Influenza viruses (H1N1, H7N1 and H7N9), human immunodeficiency virus, Ebola virus, coronaviruses as SARS-CoV, MERS and the SARS-CoV-2 causing the COVID-19 pandemic have been the most significant.

SARS-CoV-2 was identified in China at the end of 2019 [11] and has become the first pandemic coronavirus (CoV). After the first case report of the presence of SARS-CoV-2 RNA in faeces [12], and because of the presence in the past of SARS-CoV-1 in faeces and sewage [13–15], the scientific community started to investigate if this virus could spread into the environment. Specific stability of SARS-CoV-2 has only been tested in aerosols and surfaces [16], but it is known that enveloped viruses are capable of retaining infectivity for days to months in aqueous environments [17–19]. On



Evaluation of two rapid ultrafiltration-based methods for SARS-CoV-2 concentration from wastewater

E. Forés^{a,b}, S. Bofill-Mas^{a,b}, M. Itarte^{a,b}, S. Martínez-Puchol^{a,b}, A. Hundesa^a, M. Calvo^c, C.M. Borrego^{d,e}, L.L. Corominas^{d,f}, R. Girones^{a,b}, M. Rusiñol^{g,*}

^a Section of Microbiology, Virology and Biotechnology, Department of Genetics, Microbiology and Statistics, University of Barcelona, Spain

^b The Water Institute of the University of Barcelona, Spain

^c Section of Statistics, Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, Spain

^d Catalan Institute for Water Research (ICRA), Scientific and Technological Park of the University of Girona, Emili Grahit 101, E-17003 Girona, Spain

^e Group of Molecular Microbial Ecology, Institute of Aquatic Ecology, University of Girona, E-17003 Girona, Spain

^f University of Girona, E-17003 Girona, Spain

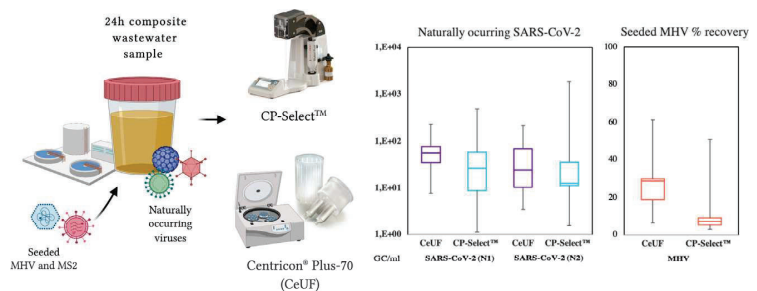
^g Institute of Environmental Assessment & Water Research (IDAEA), CSIC, Barcelona, Spain



HIGHLIGHTS

- Centricon® and CP-Select™ performed equally for naturally occurring SARS-CoV-2
- Higher MHV recoveries were calculated using centrifugal ultrafiltration devices.
- Naturally occurring viruses complement concentration methods comparison.
- A 23% of detected SARS-CoV-2 adsorb to the solid fraction of wastewater.
- CP-Select™ fits into a BSL-2 cabinet enabling to work under biosafety containment.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 1 November 2020

Received in revised form 21 December 2020

Accepted 21 December 2020

Available online 7 January 2021

Editor: Damia Barcelo

Keywords:

SARS-CoV-2

Wastewater

Viral concentration method

Ultrafiltration

Viral recovery

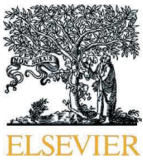
ABSTRACT

Quantitative measurements of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in raw wastewater have been implemented worldwide since the beginning of the pandemic. Recent efforts are being made to evaluate different viral concentration methodologies to overcome supplier shortages during lockdowns. A set of 22-wastewater samples seeded with murine hepatitis virus (MHV), a member of the *Coronaviridae* family, and the bacteriophage MS2, were used to characterize and compare two ultrafiltration-based methods: a centrifugal ultrafiltration device (Centricon® Plus-70) and the automated concentrating pipette CP-Select™. Based on the recovery efficiencies, significant differences were observed for MHV, with Centricon® Plus-70 (24%) being the most efficient method. Nevertheless, concentrations of naturally occurring SARS-CoV-2, Human adenoviruses and JC polyomaviruses in these samples did not result in significant differences between methods suggesting that testing naturally occurring viruses may complement the evaluation of viral concentration methodologies. Based on the virus adsorption to solids and the necessity of a pre-centrifugation step to remove larger particles and avoid clogging when using ultrafiltration methods, we assessed the percentage of viruses not quantified after ultrafiltration. Around 23% of the detected SARS-CoV-2 would be discarded during the debris removal step. The CP-Select™ provided the highest concentration factor (up to 333×) and the lowest LoD (6.19×10^3 GC/l) for MHV and proved to be fast, automatic, highly reproducible and suitable to work under BSL-2 measures.

© 2021 Elsevier B.V. All rights reserved.

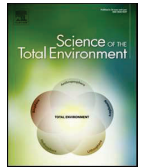
* Corresponding author.

E-mail address: marta.rosinol@idaea.csic.es (M. Rusiñol).



Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Monitoring waves of the COVID-19 pandemic: Inferences from WWTPs of different sizes



M. Rusiñol^{a,1}, I. Zammit^{b,c,1}, M. Itarte^d, E. Forés^{d,e}, S. Martínez-Puchol^{d,e}, R. Girones^{d,e}, C. Borrego^{b,c}, Ll. Corominas^{b,c,*}, S. Bofill-Mas^{d,e,**}

^a Institute of Environmental Assessment & Water Research (IDAEA), CSIC, Barcelona, Spain

^b Catalan Institute for Water Research (ICRA), Emili Grahit 101, 17003 Girona, Spain

^c Universitat de Girona, Plaça de Sant Domènec 3, 17004 Girona, Spain

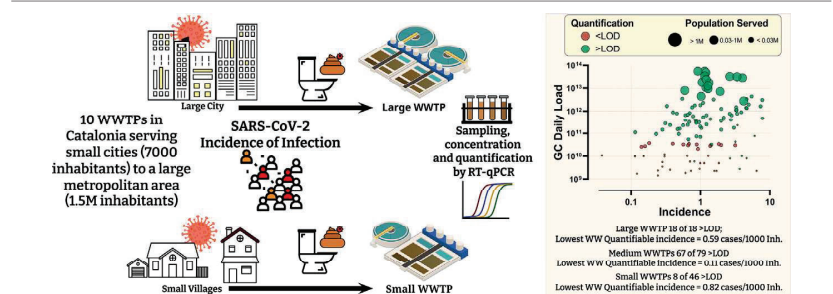
^d University of Barcelona, Section of Microbiology, Virology and Biotechnology, Department of Genetics, Microbiology and Statistics, Faculty of Biology, Barcelona, Spain

^e The Water Institute of the University of Barcelona, Spain

HIGHLIGHTS

- SARS-CoV-2 loads correlate much better than concentrations with infection incidence.
- Higher SARS-CoV-2 loads were measured during the first wave than the second wave.
- Large WWTPs have quantifiable levels at lower infection incidence than small WWTPs.
- Monitoring of small WWTPs is challenging and depends on high COVID-19 incidence.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 22 March 2021

Received in revised form 27 April 2021

Accepted 27 April 2021

Available online 7 May 2021

Editor: Damia Barcelo

Keywords:

Wastewater-based epidemiology
SARS-CoV-2
Pandemic surveillance
Small large WWTPs
Human adenoviruses
JC polyomavirus

ABSTRACT

Wastewater based epidemiology was employed to track the spread of SARS-CoV-2 within the sewershed areas of 10 wastewater treatment plants (WWTPs) in Catalonia, Spain. A total of 185 WWTPs inflow samples were collected over the period consisting of both the first wave (mid-March to June) and the second wave (July to November). Concentrations of SARS-CoV-2 RNA (N1 and N2 assays) were quantified in these wastewaters as well as those of Human adenoviruses (HAdV) and JC polyomavirus (JCPyV), as indicators of human faecal contamination. SARS-CoV-2 N gene daily loads strongly correlated with the number of cases diagnosed one week after sampling i.e. wastewater levels were a good predictor of cases to be diagnosed in the immediate future. The conditions present at small WWTPs relative to larger WWTPs influence the ability to follow the pandemic. Small WWTPs (<24,000 inhabitants) had lower median loads of SARS-CoV-2 despite similar incidence of infection within the municipalities served by the different WWTP (but not lower loads of HAdV and JCPyV). The lowest incidence resulting in quantifiable SARS-CoV-2 concentration in wastewater differed between WWTP sizes, being 0.11 and 0.82 cases/1000 inhabitants for the large and small sized WWTP respectively.

© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Correspondence to: L. Corominas, Universitat de Girona, Plaça de Sant Domènec 3, 17004 Girona, Spain.

** Correspondence to: S. Bofill-Mas, The Water Institute of the University of Barcelona, Spain.

E-mail addresses: lcorminas@icra.cat (L. Corominas), sbofill@ub.edu (S. Bofill-Mas).

¹ These authors contributed equally to this work and share first authorship.

1. Introduction

The COVID-19 pandemic, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), went from a local issue in China to a truly global threat to human health, wellbeing and the



Evaluation of a virus concentration method based on ultrafiltration and wet foam elution for studying viruses from large-volume water samples

Eva Forés^{a,b}, Marta Rusiñol^{a,c,*}, Marta Itarte^{a,b}, Sandra Martínez-Puchol^{a,b}, Miquel Calvo^a, Sílvia Bofill-Mas^{a,b}

^a Laboratory of Viruses Contaminants of Water and Food, Genetics, Microbiology & Statistics Department at the University of Barcelona (UB), Barcelona, Catalonia, Spain

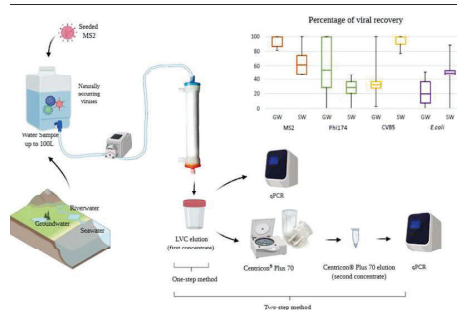
^b The Water Research Institute (IdRA), Universitat de Barcelona, Barcelona, Catalonia, Spain

^c Institute of Environmental Assessment & Water Research (IDAEA), CSIC, Barcelona, Catalonia, Spain

HIGHLIGHTS

- LVC kit is a rapid method for viral detection in water samples up to 100 L.
- The novel wet foam elution technology allows viral elution in the field.
- Method sensitivity can be increased by coupling a secondary concentration step.
- HAdV, NoV GI and NoV GII have been detected in freshwater and sea water.
- Cell culture and viral NGS studies can be conducted with the LVC eluates.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 21 December 2021

Received in revised form 4 March 2022

Accepted 5 March 2022

Available online 9 March 2022

Editor: Warish Ahmed

Keywords:

Viral detection
Viral metagenomics
Large volume concentration
Freshwater
Sea water
Ultrafiltration

ABSTRACT

Assessing the presence of viruses in large-volume samples involves cumbersome methods that require specialized training and laboratory equipment. In this study, a large volume concentration (LVC) method, based on dead-end ultrafiltration (DEUF) and Wet Foam Elution™ technology, was evaluated in different type of waters and different microorganisms. Its recovery efficiency was evaluated through different techniques (infectivity assays and molecular detection) by spiking different viral surrogates (bacteriophages PhiX174 and MS2 and Coxsackie virus B5 (CVB5) and *Escherichia coli* (*E. coli*)). Furthermore, the application of a secondary concentration step was evaluated and compared with skimmed milk flocculation. Viruses present in river water, seawater and groundwater samples were concentrated by applying LVC method and a centrifugal ultrafiltration device (CeUF), as a secondary concentration step and quantified with specific qPCR Human adenoviruses (HAdV) and noroviruses (NoVs). MS2 was used as process control, obtaining a mean viral recovery of $22.0 \pm 12.47\%$. The presence of other viruses was also characterized by applying two different next-generation sequencing approaches. LVC coupled to a secondary concentration step based on CeUF allowed to detect naturally occurring viruses such as HAdV and NoVs in different water matrices. Using HAdV as a human fecal indicator, the highest viral pollution was found in river water samples (100% of positive samples), followed by seawater (83.33%) and groundwater samples (66.67%). The LVC method has also proven to be useful as a virus concentration method in the field since HAdV and NoVs were detected in the river water and groundwater samples concentrated in the field. All in all, LVC method presents high concentration factor and a low limit of detection and provides viral concentrates useful for subsequent molecular analysis such as PCR and massive sequencing.

* Corresponding author at: Laboratory of Viruses Contaminants of Water and Food, Genetics, Microbiology & Statistics Department at the University of Barcelona (UB), Barcelona, Catalonia, Spain.

E-mail address: mrusinol@ub.edu (M. Rusiñol).



OPEN

The Catalan Surveillance Network of SARS-CoV-2 in Sewage: design, implementation, and performance

Laura Guerrero-Latorre^{1,2,14}, Neus Collado^{1,2,14}, Nerea Abasolo³, Gabriel Anzaldí⁴, Sílvia Bofill-Mas⁵, Albert Bosch^{6,7}, Lluís Bosch^{1,2}, Sílvia Busquets^{1,2}, Antoni Caimari⁸, Núria Canela³, Albert Carcereny^{6,7}, Carme Chacón⁹, Pilar Ciruela⁹, Irene Corbella⁹, Xavier Domingo⁴, Xavier Escoté¹⁰, Yaimara Espiñeira⁴, Eva Forés⁵, Isabel Gandullo-Sarró¹¹, David Garcia-Pedemonte^{6,7}, Rosina Girones⁵, Susana Guix^{6,7}, Ayalkibet Hundesa⁵, Marta Itarte⁵, Roger Mariné-Casadó¹⁰, Anna Martínez⁹, Sandra Martínez-Puchol⁵, Anna Mas-Capdevila¹⁰, Cristina Mejías-Molina⁵, Marc Moliner i Rafa¹¹, Antoni Munné¹¹, Rosa Maria Pintó^{6,7}, Josep Pueyo-Ros^{1,2}, Jordi Robusté-Cartró¹¹, Marta Rusiñol¹², Robert Sanfeliu⁴, Joan Teichenné¹⁰, Helena Torrell³, Lluís Corominas^{1,2} & Carles M. Borrego^{1,13}✉

Wastewater-based epidemiology has shown to be an efficient tool to track the circulation of SARS-CoV-2 in communities assisted by wastewater treatment plants (WWTPs). The challenge comes when this approach is employed to help Health authorities in their decision-making. Here, we describe the roadmap for the design and deployment of SARSAIGUA, the Catalan Surveillance Network of SARS-CoV-2 in Sewage. The network monitors, weekly or biweekly, 56 WWTPs evenly distributed across the territory and serving 6 M inhabitants (80% of the Catalan population). Each week, samples from 45 WWTPs are collected, analyzed, results reported to Health authorities, and finally published within less than 72 h in an online dashboard (<https://sarsaigua.icra.cat>). After 20 months of monitoring (July 20–March 22), the standardized viral load (gene copies/day) in all the WWTPs monitored fairly matched the cumulative number of COVID-19 cases along the successive pandemic waves, showing a good fit with the diagnosed cases in the served municipalities (Spearman Rho = 0.69). Here we describe the roadmap of the design and deployment of SARSAIGUA while providing several open-access tools for the management and visualization of the surveillance data.

Shortly after the onset of the outbreak of the current COVID-19 pandemic, scientists renewed their interest in the application of wastewater-based epidemiology (WBE)¹ to track the communal circulation of SARS-CoV-2 through the quantification of its genetic traces in sewage². This resulted in an overwhelming number of

¹Catalan Institute for Water Research (ICRA), Emili Grahit 101, 17003 Girona, Catalonia, Spain. ²Universitat de Girona, 17003 Girona, Catalonia, Spain. ³Centre for Omic Sciences (COS), Joint Unit Universitat Rovira i Virgili-EURECAT, Unique Scientific and Technical Infrastructures (ICTS), Eurecat, Centre Tecnològic de Catalunya, Avinguda Universitat 1, 43204 Reus, Catalonia, Spain. ⁴Applied Artificial Intelligence Technological Unit, Eurecat, Centre Tecnològic de Catalunya, Science and Technology Park, H3, 25003 Lleida, Catalonia, Spain. ⁵Laboratory of Viruses Contaminants of Water and Food, Department of Genetics, Microbiology and Statistics, School of Biology, University of Barcelona, 08028 Barcelona, Catalonia, Spain. ⁶Enteric Virus Laboratory, Section of Microbiology, Virology and Biotechnology, Department of Genetics, Microbiology and Statistics, School of Biology, University of Barcelona, 08028 Barcelona, Catalonia, Spain. ⁷Research Institute of Nutrition and Food Safety (INSA), University of Barcelona, 08921 Santa Coloma de Gramenet, Catalonia, Spain. ⁸Biotechnology Area, Eurecat, Centre Tecnològic de Catalunya, Avinguda Universitat 1, 43204 Reus, Catalonia, Spain. ⁹Public Health Agency of Catalonia (ASPCAT), Roc Boronat 81-95, 08005 Barcelona, Catalonia, Spain. ¹⁰Technological Unit of Nutrition and Health, Eurecat, Centre Tecnològic de Catalunya, Avinguda Universitat 1, 43204 Reus, Catalonia, Spain. ¹¹Catalan Water Agency (ACA), Provença, 260, 08008 Barcelona, Catalonia, Spain. ¹²Institute of Environmental Assessment & Water Research (IDAEA), CSIC, Barcelona, Catalonia, Spain. ¹³Group of Molecular Microbial Ecology, Institute of Aquatic Ecology, University of Girona, 17003 Girona, Catalonia, Spain. ¹⁴These authors contributed equally: Laura Guerrero-Latorre and Neus Collado. ✉email: cborrego@icra.cat



Evaluation of pathogen disinfection efficiency of electrochemical advanced oxidation to become a sustainable technology for water reuse

Eva Forés ^{a,b}, Cristina Mejías-Molina ^{a,b}, Arantxa Ramos ^c, Marta Itarte ^{a,b}, Ayalkibet Hundesa ^{a,b}, Marta Rusiñol ^{a,b}, Sandra Martínez-Puchol ^{a,b}, Pau Esteve-Bricullé ^c, Alejandro Espejo-Valverde ^c, Ignasi Sirés ^d, Miquel Calvo ^e, Rosa M. Araujo ^c, Rosina Girones ^{a,b,*}

^a Laboratory of Viruses Contaminants of Water and Food, Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona (UB), Barcelona, Spain

^b Institut de Recerca de l'Aigua (IdRA), Universitat de Barcelona (UB), Barcelona, Spain

^c Secció de Microbiologia, Virologia i Biotecnologia, Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona (UB), Barcelona, Spain

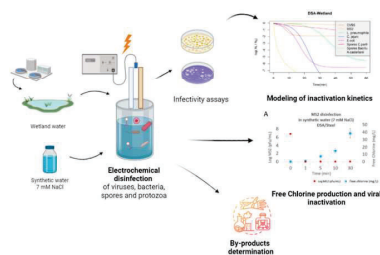
^d Laboratori d'Electroquímica dels Materials i del Medi Ambient, Departament de Ciència de Materials i Química Física, Secció de Química Física, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1-11, 08028, Barcelona, Spain

^e Secció d'Estadística, Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona (UB), Barcelona, Spain

HIGHLIGHTS

- Electrochemical disinfection for safe water production with low energy consumption.
- Disinfection of bacteria, bacteria spores, bacteriophages, virus, and protozoa.
- Description of inactivation kinetics using logistic, exponential, and linear models.
- Disinfection efficiency is limited for virus, spores and amoebas.
- Bacteria and bacteriophages are not a suitable model for viral disinfection.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:

Electrochemical disinfection
Microbial pathogens
Microbial indicators
Inactivation kinetics
Viruses
Water reuse

ABSTRACT

Water treatment and reuse is gaining acceptance as a strategy to fight against water contamination and scarcity, but it usually requires complex treatments to ensure safety. Consequently, the electrochemical advanced processes have emerged as an effective alternative for water remediation. The main objective here is to perform a systematic study that quantifies the efficiency of a laboratory-scale electrochemical system to inactivate bacteria, bacterial spores, protozoa, bacteriophages and viruses in synthetic water, as well as in urban wastewater once treated in a wetland for reuse in irrigation. A Ti/RuO₂-based plate and Si/BDD thin-film were comparatively employed as the anode, which was combined with a stainless-steel cathode in an undivided cell operating at 12 V. Despite the low resulting current density (<15 mA/cm²), both anodes demonstrated the production of oxidants in wetland effluent water. The disinfection efficiency was high for the bacteriophage MS2 (T99 in less than 7.1 min) and bacteria (T99 in about 30 min as maximum), but limited for CBV5 and TuV, spores and amoebas (T99 in more than 300 min). MS2 presented a rapid exponential inactivation regardless of the anode and bacteria

* Corresponding author. Laboratory of Viruses Contaminants of Water and Food, Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona (UB), Barcelona, Spain.

E-mail address: rgirones@ub.edu (R. Girones).

<https://doi.org/10.1016/j.chemosphere.2022.137393>

Received 22 September 2022; Received in revised form 4 November 2022; Accepted 24 November 2022

Available online 26 November 2022

0045-6535/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

RESEARCH

Open Access



Wastewater-based epidemiology applied at the building-level reveals distinct virome profiles based on the age of the contributing individuals

Cristina Mejías-Molina^{1,2*}, Anna Pico-Tomàs³, Sandra Martínez-Puchol¹, Marta Itarte^{1,2}, Helena Torrell⁴, Núria Canela⁴, Carles M. Borrego^{3,5}, Lluís Corominas³, Marta Rusiñol^{1,2} and Sílvia Bofill-Mas^{1,2}

Abstract

Background Human viruses released into the environment can be detected and characterized in wastewater. The study of wastewater virome offers a consolidated perspective on the circulation of viruses within a population. Because the occurrence and severity of viral infections can vary across a person's lifetime, studying the virome in wastewater samples contributed by various demographic segments can provide valuable insights into the prevalence of viral infections within these segments. In our study, targeted enrichment sequencing was employed to characterize the human virome in wastewater at a building-level scale. This was accomplished through passive sampling of wastewater in schools, university settings, and nursing homes in two cities in Catalonia. Additionally, sewage from a large urban wastewater treatment plant was analysed to serve as a reference for examining the collective excreted human virome.

Results The virome obtained from influent wastewater treatment plant samples showcased the combined viral presence from individuals of varying ages, with astroviruses and human bocaviruses being the most prevalent, followed by human adenoviruses, polyomaviruses, and papillomaviruses. Significant variations in the viral profiles were observed among the different types of buildings studied. Mamastrovirus 1 was predominant in school samples, salivirus and human polyomaviruses JC and BK in the university settings while nursing homes showed a more balanced distribution of viral families presenting papillomavirus and picornaviruses and, interestingly, some viruses linked to immunosuppression.

Conclusions This study shows the utility of building-level wastewater-based epidemiology as an effective tool for monitoring the presence of viruses circulating within specific age groups. It provides valuable insights for public health monitoring and epidemiological studies.

Keywords Human age-related virome, Building-level, Passive samplers, Wastewater-based epidemiology, Targeted enrichment sequencing

*Correspondence:

Cristina Mejías-Molina
cmejias@ub.edu

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

