UNIVERSITAT DE BARCELONA

Aplicació de tècniques de seqüenciació massiva a l´estudi de virus potencialment contaminants d´aigües i/o aliments

Xavier Fernández Cassi

Aplicació de tècniques de seqüenciació massiva a l'estudi de virus potencialment contaminants d'aigües i/o aliments

Per

Xavier Fernández Cassi

Maig 2017

UNIVERSITAT DE BARCELONA

Programa de doctorat:
Microbiologia Ambiental i Biotecnologia
Departament de Genètica,
Microbiologia i Estadística

TESI DOCTORAL

Aplicació de tècniques de seqüenciació massiva a l'estudi de virus potencialment contaminants d'aigües i/o aliments

Memòria presentada per

Xavier Fernández Cassi

Per optar al grau de

Doctor per la Universitat de Barcelona

Tesi realitzada sota la direcció de la Dra. Rosina Girones Llop i la Dra. Sílvia Bofill Mas,

Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona.

Co-directora i tutora	Co-directora	Autor
Rosina Girones Llop	Sílvia Bofill Mas	Xavier Fernández Cassi

A LA MEVA FAMÍLIA I AMICS

AGRAÏMENTS

La tesi és com una "*muntanya russa de les emocions*" amb molt poques pujades i unes baixades dures i impactants. En aquest camí he pogut establir vincles amb les persones i personatges que circulen per la realitat paral·lela que representa el Departament de Microbiologia (nomenat ara, tècnicament, secció de Microbiologia). Vull donar-vos les gràcies a tots aquells que heu estat acompanyant-me especialment durant les baixades.

Voldria agrair, en primer lloc, a la Laura Guerrero ja que és la responsable que estigui ara escrivint aquestes paraules; sense aquella conversa amb ella enmig del Departament (quan la UB era tota una novetat), no hi hauria tesi. Qui ens hauria de dir que dues de les "Black Sheep" de la UPF acabarien al mateix grup fent el doctorat? Si ens veies el Pérez... FLIPARIA! Aigua i oli que es barregen… Així és com ens definiria. Moltes gràcies per haver-me deixat entrar en el teu cercle d'amistats i gràcies per haver-me fet conèixer la Marta Rusiñol també coneguda com a *gadella del meu cor*, *Rusi*, *Martus*, *Marti* o, científicament parlant, com a *Marta Rusialt165ol*. Amb tu HE VISCUT la tesi i totes les seves pujades i baixades. Gràcies per ensenyar-me TOT el que sé sobre mostrejos, flocular, concentrar, extraure (encara que fos a la poiata) i fer qPCR. A costa teva he après que realment és possible treballar molt sense renunciar a un bon dinar. Pels moments viscuts de trivial, de cotxe (amb multa inclosa), de riures, de distraccions (la majoria teves sempre, vegetal!), algun que altre *bailoteo* i pels moments que ens queden per viure. No sé que en treuré de tot això de la tesi, però de ben segur una AMIGA per sempre! Saps escollir molt bé; això va per tu Alfio: *l'olio di Olesa è meglio di quello di Foggia*! Cuideu molt de la vostra petita Anna, té les qualitats musicals del tiet… pot arribar molt lluny!

A mi chica, Sandra Martínez-Puchol, coneguda com a *Trans*. És fort que amb un any i poc que fa que ens coneixem des del minut 0 ens duguem tan bé; ha estat un "*flechazo*"! Gràcies per aparèixer quan més et necessitava (hi havia una baixada molt forta). No decaiguis mai, tens un esperit inquiet que *"¡ya les gustaria a muchos!"*. M'hauria agradat poder compartir molts més moments de tesi al teu costat, però tranquil·la... *¡no te librarás de mi tan fácilmente!* 😝. Gràcies també per tot el suport logístic durant aquest *sprint* final! *I own you one*!

Als nens més guapos que ha tingut Vircont, això va per vosaltres Eva i David (rateta). Sou un sol i valeu moltíssim. David moltes gràcies per estar sempre amb voluntat d'ajudar i no tenir un NO per resposta, això si…, encén la Xbox per jugar😝. Ets la meva rata de laboratori japonesa preferida! Eva, gràcies per proveir-me de cafè quan l'he necessitat i agafar patates fregides encara que no t'agradin! El teu humor àcid i els teus "*zascas*" m'encanten! No canvis mai, ets molt autèntica Piscis!

Natalia, alies la tecles (o tecles-radial si estàs a la planta 0), encara que no has estat físicament en molts moments viscuts, sempre has estat present telemàticament. Aquesta tesi, sense tu, simplement no existiria. Hem fet junts un camí, no sense problemes, així que moltes gràcies per tenir tanta paciència i ves per on… al final acabarem sincronitzats! M'agradaria també nomenar i agrair al teu pare adoptiu, en Josep, per tot el suport rebut. Persia, la meva veritable deixeble, a la que sempre recordo amb un somriure, merci pels moments que m'has donat. Marta Itarte voldria desitjar-te molta sort, tens molt potencial! Pete you are a cool guy, but you need to stop eating chicken burguers! Texenery, no me olvido de ti tampoco, echo de menos tus dos litros de colonia! Hagar, you left your mark in all VIRCONT members, hope to see you soon, my speak English, b****, It's America! Agrair també a l'Ayalke i l'Eloy per la paciència que han tingut amb mi; sé que no sóc fàcil.

Tot i no ser tècnicament de VIRCONT, hem compartit molt de temps junts amb els Araujistes o Araujitos. Tinc un especial record de la Silvia Cervero (Cerver) amb la que hem passat unes grans estones durant la primera etapa d'aquesta tesi. Esperar un bus sense tu a la diagonal NO és el mateix. També agrair a la Carmina (Karmins) tot el bé que ens ha cuidat sempre i per donar-me la seva fantàstica companyia durant les firmes o concerts de Fangoria, quin parell! Moltes gràcies per posar els accents on tocava, *jajaja*. També a l'Edu i als seus shows; tot un personatge que es fa estimar... Ara sí que s'acaba el *reality*, titi: *estamos expulsados*! Yexe y Alberto gracias por intentar siempre apaciguar las aguas y traer calma al laboratorio 8. Gràcies també a tota l'altra gent del laboratori 8 que ja no hi són siguin Araujitos/Araujistes o Rosinots/Honeys. De tots he après alguna cosa: Byron, Jesús, Anna, Sandra, Tarik, Idaira, Marcelle, Cagla, Sara i Toni (petit xef).

A la gent del lab 5, la meva segona casa després del "destierro" del lab 8, gràcies per rebre'm amb els braços oberts. Especialment voldria agrair-li a la Sara (petarda) les vegades que m'ha fet riure; m'encantaves com a alumne i m'has enamorat com a companya. També voldria agrair al David (Madonno) els moments de riures junts… Recordes el...moment "ve un Cubano"? (*wink, wink*). Als dos nens fantàstics que tenen al 5 que són un amor. Sí, vosaltres en Miquel (teòricament és ballarí però ningú l'ha vist mai ballant res) i la Cristina (tsundere) que sempre em fa riure amb les seves teories (com la de Anaconda és un videoclip feminista, *me parto*). Sabrina, Melania mola más que Juliana. Marga, una pena no haver co-habitat al lab! Magda, m'ha agradat molt compartir laboratori amb vosaltres i poder participar dels vostres càstings, tota una experiència!

A les meves noies del lab 3-6. Sí, sí, vosaltres, l'ENEMIC *hahah! ¡Me parto!* Amb vosaltres he viscut alguns dels millors moments de la tesi: findes, menjars de morses, congressos "Nacionales" XD, diades per la independència (Catalunya

lliure!) i viatges exòtics (esa Thailandia!). Montse (Montserrateta), gràcies per fer-me sempre costat i ser la meva no-compi de gym però sí de pilates! *T'estim* amb bogeria! M'has ajudat moltíssim i m'has escoltat molt més! Molta força! HO TENS! Anna, la meva arpia voladora, gràcies per tots els moments i perdona que et fes el spoiler de la mort de Ned... Pensava que ho sabies... Aurora o Auroreta gràcies per ser com ets, per les tornades a casa parlant com frikis dels *noros*, *sapos* i *astros*, però també pels *bailoteos* i els riures, ets un sol! Cecilia, mi Argentina/Americana favorita, em va encantar coneixe't i poder disfrutar dels "*jalus gourmet*" i el sofriment a la UPC.

Voldria agrair també a tota la gent dels labs 9-10, que tot i el nom d'extraterrestre (MARS), són bona gent. Aisss esa Laura Sala (sempre has volgut ser VIRCONT, *eh granuja, lo sabes!*), la Míriam Pasqual (sempre tan *sporty* ella fent coses rares com el "*Touch*"), i la Júlia (Ánimo rajada :P), Pablo (tu también necesitas algo de ánimo; venga, el último *sprint!*) y Maryuri (gracias por salvarme siempre que se me cierra la puerta de los labavos). Rachi eres un amor, y por eso te perdono aunque me hayas dejado tirado en el Diablo y nunca me hayas invitado a jugar al *Smash*... las debilitades son debilidades, no me puedo enfadar contigo. Aiora gràcies per soportar les meves intromissions en les teves coses amb tan de temple, i ajudar-me amb les cèl·lules i les comandes! (mami que manos tienes, visítame más!). A la jefa de tota aquesta *troupe*, Maite em fa il·lusió/por que siguis el meu tribunal, espero no decebre't! Voldria també mencionar a tota la gent d'aquests dos labs que ja no hi és: l'Alex (ladilla que pesado que eras y cuánto te hemos echado de menos), Miriam alias Parru (usurpadora, *enxufam* ja tia!), Maru (enquincla babosa!!), Arnau i Martes. No voldria oblidar-me de la meva postdoc preferida l'Elisenda; ets un solete. Ni a la Sonia Paytubi (mi *partner* sin lactosa) ni al seu clon (Estel, ets la companya de

departament que més hauria volgut, i *"¡si no lo digo reviento!"*). Pedrito, tu eres toda una revelación! Molta *energy*!

No em voldria deixar a les meves noies de secretaria que sempre saben aconsellar-me i m'ajuden a canalitzar la fera que porto dins; gràcies Bea (Beatriu), Susana, Rosario (Charito), Mònica i Maku. No em voldria deixar tampoc l'agraïment al Manolo per tots els cops que m'ha hagut d'explicar com va el procés de la tesi. *Esto va para todos... que miembro os sacará de fiesta por "el arena"* *ahora?!*
Voldria recordar a la gent que vaig conèixer al principi de la tesi a Farmàcia. En aquell moment en que no les tenia totes de venir a Bio; si tenia dubtes és perquè sou meravelloses. Maria Jesús (*la mujer antientropia, si tuviera una empresa te contrataba*), la Carme (sempre esperant amb un dolç somriure), la Carla (cochi I Love U!), Ornella (Perri), Eleonora me echabas tanto de menos que viniste a Bio, i la Núria (Perri Maison).

I would like also to thank my lovely people at DTU: Mathis, Anne and Anna Charlotte. Thanks for being so nice while I was there. You made me feel like home. Specially, I would like to say thanks to my Danish/Sweedish girls Kat and Maria. It was awesome to stay with you in DK, BCN and TOKYO!

A tots els amics que he pogut fer "treballant" de congrés o de cursillo: Carmen, Noelia, Miguel Ángel "el canario", Miguel, Ion, Angel, Narciso, Quique i Irene.

Als meus amics de tota la vida... els que em soporteu de sempre, GRÀCIES. Pablo, hermano, ladilla y fístula duodenal. Gracias por estar ahí siempre (incluso cuando me dejo el pasaporte) Todas esas chuches financiadas te las he devuelto bien, ¿eh?... presentándote a Marta Martí (alias Lauren Bennet) menudo pivón de mujer y madre te he conseguido! Marta, gràcies per fer-te

estimar d'una forma tan natural, és com si et conegués des de fa mil!! Ganas corazones igual de bien que vendes las motos 😉.

Martilla... Gracias por ofrecerme siempre tu ayuda y escuchar todas mis ralladas (especialmente las profesionales), y las que te quedan por soportar porque nunca te librarás de mí, *hahaha*. A mis amigos fiesteros: Miquel, Fran y Miguel. Gràcies per treure'm de passeig de tant en tant i donar-me converses mundanas nivell màxim! Pumpy gràcies per arrencar-me sempre uns quants somriures, eres mi "simple diplomada" favorita. A les meves amigues biòlogues Albins, Maki i Pinxus gràcies per estar encara formant part de la meva vida, des que vam decidir iniciar aquesta "carrera junts" a la Pompeu (recordeu, som l'elit, JUAS). Ara ja som tots doctors... quan ho petem fent alguna col·laboració? Gràcies també al Ferran! Sense tu aquesta portada tan surrealista no seria possible. A les meves noies de CTA (Marian, Esther, Lorena, Gemma, Marta Martí (again) i Patri) gràcies per donar-me la millor etapa universitària que hauria pogut somiar.

M'agradaria també agrair a la meva família el suport que m'han demostrat al llarg d'aquests anys. Potser algunes vegades no han entès el que significa fer una tesi i més d'una vegada ha aparegut la recurrent pregunta de: quant començaràs a treballar? O perquè no et busques una feina on se't valori? Al llarg dels anys han entès que això no es fa per diners... sinó perquè tenen un fill/germà una mica *friky*. Gràcies a la meva mare Antònia per ser (simplement) la millor mare del món, encara que mai entengui res del que em passa i al meu pare Enrique encara que sigui un cas! Gràcies també a les meves germanes i germà Elvi, Isa, Amparo i Josemi per cuidar-me sempre com un fill més, ja sabeu... "el nen". Gràcies a les meves nebodes i nebots (Elvira, Mar, Mireia, Helena, Marina J, Pau, Marina F i Marc) per ser tan guerrers i fer-me sentir el tiet guay. No voldria deixar-me els meus cunyats i cunyada Lluís, Miquel i Vicente i Marga que sempre estan al cantó disposats a ajudar.

A tots els que durant aquesta etapa han conviscut amb mi als diferents pisos: Tere, Juanmi, Tánia (peluquera), Craig (vago, trabaja y deja el vino!) i Edu. En especial als meus actuals companys de pis perquè son increïbles: Javi, Elba i Naty.

Finalment m'agradaria agrair la tesis a les meves directores: Silvia i Rosina. Sou nit i dia, llum i foscor, blanc i negre, sal i sucre i he après de tots els vostres contrastos. Moltes gràcies per donar-me l'oportunitat de gaudir d'aquesta experiència durant la tesi al vostre costat. Silvia gràcies per donar-me una relació tan propera i oferir-me la teva amistat. Encara recordo aquella conversa al bar de geo (q ni existeix ja!) en la que em vas oferir l'oportunitat d'entrar a VIRCONT. Gràcies per creure i apostar per mi. Rosina de tu he après que els dies tenen més de 24h. Al final tampoc no ha anat tan malament, no? Gràcies per donar-me l'oportunitat d'expressar-me sempre i per ensenyar-me ciència des de tots els prismes possibles, la ciència és molt més que estar assegut pipetejant.

Si has llegit això i estàs decebut perquè no apareix el teu nom, HO SENTO! No és personal!! Si no em coneixes de res i estàs dubtant en començar la tesi, un consell: HUYE!!! 😉

CONTINGUTS

ABREVIACIONS

AAstV	*Avastrovirus*
ACA	Agència Catalana de l'Aigua
ADN	Àcid desoxiribonucleic
AdV	Adenovirus
AN	Àcid nucleic
ARN	Àcid ribonucleic
AstV	Astrovirus
BKPyV	Poliomavirus BK
bp	Parell de bases (de l'anglès "base pair")
CC	Càncer de Cèrvix
CDC	Centre de control de malalties (de l'anglès "Centers for Disease Control")
cDNA	ADN de cadena complementària
CEE	Comunitat Econòmica Europea
CFU	Unitats formadores de colònies. En anglès "Colony Forming Units"
CPE	Efecte citopàtic (de l'anglès "Cytopathic Effect")
DBO	Demanda biològica d'oxigen
dsDNA	ADN bicatenari (de l'anglès "double strand DNA")
dsRNA	ARN bicatenari (de l'anglès "double strand RNA")
E.coli	*Escherichia coli*
EDAR	Estació Depuradora d'Aigües Residuals
EF	Enterococs fecals

EFSA	Autoritat Europea de protecció alimentària (de l'anglès "European Food Safety Authority")
EMA	Monoàcida d'etidi (de l'anglès "ethidium monoazide")
et al.	i col·laboradors (del llatí "*et alii*")
EV	Enterovirus
FAO	Organització de les nacions unides per l'Alimentació i l'Agricultura (de l'anglès "Food and Agriculture Organization")
FIB	Indicadors fecals bacterians (de l'anglès "Fecal Indicator Bacteria")
FFU	Unitats formadores de fluorescència (de l'anglès "Fluorescence Forming Units")
Gb	Gigabyte
GC	Còpies genòmiques
GG	Genogrup
HAdV	Adenovirus humà
HAV	Virus de l'hepatitis A
HBoV	Bocavirus humans
HBuV	Bufavirus humans
HCl	Àcid clorhídric
HEV	Virus de l'hepatitis E
HEPES	*4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid*
HPV	Papil·lomavirus humà
HPyV	Poliomavirus humà
HRoV	Rotavirus humà

ICC-PCR	PCR integrada amb cultiu cel·lular (de l'anglès "Integrated Cell Culture-PCR")
ICTV	Comitè internacional de taxonomia de virus (de l'anglès "International Comitee on Taxonomy of Viruses")
ISFET	Transistor d'efecte de camp sensible a ions (de l'anglès "Ion sensitive Field Effect Transistor")
ISO	Organització Internacional d'estandardització (de l'anglès "International Standard Organization")
JCPyV	Poliomavirus JC
Kb	Quilo-bases
KDa	Quilo-daltons
KIPyV	Poliomavirus KI
MAstV	*Mamastrovirus*
MCPyV	Poliomavirus de cèl·lules de Merkel
MST	Traçar la font de microorganismes (de l'anglès "Microbial Source Tracking")
mS/cm^2	Mil·li-Siemens per centímetre quadrat
MWPyV	Poliomavirus de Malawi, també anomenat MXPyV (poliomavirus de Mèxic) i HPyV10 (poliomavirus número 10)
MAF	Filtració d'afinitat monolítica (de l'anglès "Monolithic Affinity Filtration")
MDA	Amplificació de desplaçament múltiple (de l'anglès "Multiple Displacement Amplification")
MO	Matèria orgànica
NaOH	Hidròxid de sodi

NGS	Seqüenciació de segona generació (de l'anglès "Next-Generation Sequencing")
NS	Regió no estructural del genoma (de l'anglès "non structural")
NSJPyV	Poliomavirus de Nova Jersei, també anomenat poliomavirus 13
NoV	Norovirus
OMS	Organització Mundial de la Salut
ONU	Organització de les Nacions Unides
ORF	Marc obert de lectura (de l'anglès "Open Reading Frame")
PBS	Tampó buffer salí (de l'anglès "Phosphate-Buffer Saline")
PBV	Picobirnavirus
PCR	Reacció en cadena de la Polimerasa (de l'anglès "Polymerase Chain Reaction")
PCV	Circovirus porcí
PEG	Precipitació amb polietilenglicol (de l'anglès "PolyethileneGlycol")
PFU	Unitats formadores de clapes (de l'anglès "Plaque Forming Units")
PMA	Propidi monoàcida (de l'anglès "Propidium monoazide")
PML	Leucoencefalopatia Multifocal Progressiva
PV	Papil·lomavirus
PyV	Poliomavirus
qPCR	PCR quantitativa
RCA	Amplificació de Cercle Rodant (de l'anglès "Rolling Circle Amplification")
RD	Reial Decret

RV	Rotavirus
SaV	Sapovirus
SISPA	Amplificació independent de seqüència basada en encebador (de l'anglès "Sequence-Independent single-primer amplification")
ssDNA	ADN monocatenari (de l'anglès "single strand DNA")
ssRNA	ARN monocatenari (de l'anglès "single strand RNA")
spp.	Espècie
STLPyV	Poliomavirus de Sant Louis
SV40	Virus de simi 40
T-Ag	Antigen tumoral gran
t-Ag	Antigen tumoral petit
$TCID_{50}$	Quantitat de virus que infecten el 50% de les cèl·lules d'un cultiu (de l'anglès "50% tissue culture infective dose")
TFF	Filtració i ultrafiltració tangencial (de l'anglès "Tangential Flow Filtration")
TSPyV	Poliomavirus de la Tricodisplàsia espinulosa
UE	Unió Europea
USEPA	Agència de protecció Ambiental dels Estats Units (de l'anglès "United States Environmental Protection Agency")
UV	Llum Ultraviolada
VFF	Filtració en vòrtex (de l'anglès "Vortex Flow Filtration")
VP	Proteïna de la càpsida viral (de l'anglès "Virion Protein")
VIRADEL	Mètodes de concentració de virus basats en adsorció-elució (de l'anglès "VIRal ADsorption and ELution")

WHIM	Síndrome de berrugues, hipogammaglobulinèmia, Inmunodeficiència i Myelokathexis (de l'anglès "Warts, Hypogammaglobulinemia, Immunodeficiency, and Myelokathexis)
WUPyV	Poliomavirus WU
X g	Unitats de força centrífuga, gravetat
ZMW	Guies de zona zero (de l'anglès "Zero-Mode Waveguides")

GLOSSARI

Amplicon sequencing o metagenòmica dirigida: Aproximació per a estudiar la variació genètica d'una regió concreta utilitzant tècniques de seqüenciació massiva. Es requereixen un parell d'encebadors de PCR per a capturar la regió específica a estudiar, permetent identificar i caracteritzar de forma eficient les seqüències amplificades.

Anotar: Procés d'identificar la localització dels gens i de les regions codificants en un genoma i determinar-ne la funció

Contig: Fragment de DNA obtingut a partir de la superposició d'un grup de *reads*.

Ensamblar: Procés per el qual diferents fragments curts de DNA (*reads*) s'uneixen per a formar seqüències més llargues amb l'objectiu de reconstruir genomes complets.

Mapar: procés que consisteix en comparar cadascun dels reads contra un genoma de referència. El procés de mapar resulta a nivell computacional menys costós que ensamblar.

Read: Conjunt de seqüències resultant després del procés de seqüenciació

INTRODUCCIÓ

1. L'aigua en l'estudi dels virus

1.1. L'aigua un recurs escàs

L'aigua és un recurs imprescindible i escàs per als essers humans tan des d'un punt de vista purament biològic com també social i econòmic. La majoria de l'aigua present al nostre planeta és salada en forma de mars i oceans. Només una petita fracció del total d'aigua del planeta (2,8%) és dolça, i per tant, apta per al seu consum. Segons estimacions de l'Organització de les Nacions Unides (ONU) la població Mundial arribarà als 9 milions d'habitants l'any 2050 (United Nations, 2013). Les projeccions actuals suggereixen que, associat a aquest creixement d'habitants, hi haurà un increment de la demanda d'aigua, a nivell mundial, del 55% entre els anys 2000 i 2050 (Gurría, 2012), evidenciant un balanç hídric insostenible a llarg termini. Per altra banda, l'efecte que el canvi climàtic pot exercir sobre les conques hídriques, disminuint-ne tant el cabal hídric com la qualitat de l'aigua, suposen un altre problema. Tot això evidencia la necessitat d'aplicar mesures urgents que permetin una reducció, reutilització i reciclatge dels recursos hídrics disponibles.

1.2. Principals agents contaminants de l'aigua

Existeixen diverses fonts de contaminació de l'aigua: residus orgànics i inorgànics generats per l'home, sediments i materials en suspensió, nutrients d'origen vegetal, microorganismes patògens, etc. Entre aquests contaminants són d'especial interès els microorganismes, i en concret, els d'origen fecal. La presència d'aquests microorganismes suposa una pèrdua de qualitat de les masses d'aigua, ja que alguns són microorganismes patògens. Segons informació extreta de l' Organització Mundial de la Salut (OMS), al voltant de 2 milions de persones moren cada any, en la seva majoria nens menors de 5 anys, degut a malalties diarreiques freqüentment associades al consum d'aigua o aliments

contaminats (WHO, 2013). Els principals patògens transmesos per aigua i aliments per via fecal-oral es presenten a la **Figura 1**.

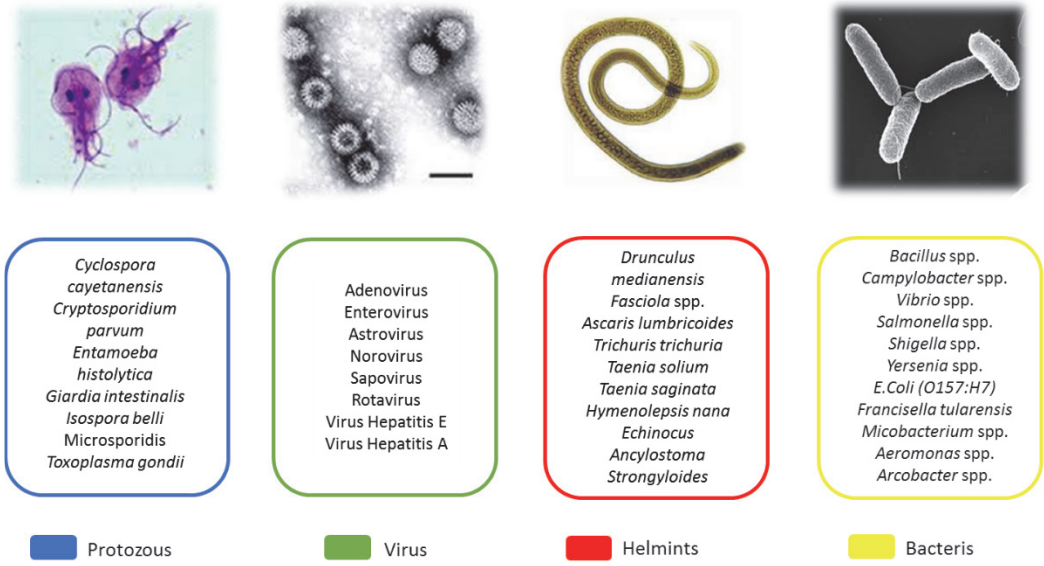


Figura 1. Principals patògens transmesos per aigua i aliments per via fecal-oral.

1.3. Els virus com a contaminants. Ús d'indicadors

Els contaminants d'origen animal i/o humà arriben a l'aigua a través de les excrecions (femta o orina) o per la descamació epitelial. Aquests contaminants s'infiltren en el sòl, ja sigui pel seu ús com a adobs o bé perquè formen part de les aigües residuals; contaminant aigües de reg, aqüífers o altres masses d'aigua **(Figura 2)**.

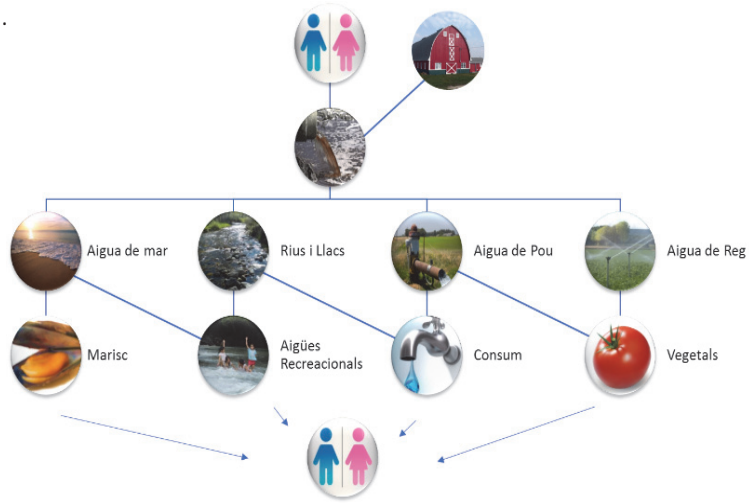


Figura 2. Fonts de disseminació de virus animals i humans a l'ambient.

La presència de microorganismes a l'aigua pot suposar un perill si aquesta entra en contacte amb aliments com vegetals, fruites, o mol·luscs bivalves, pot afectar la qualitat d'aigües recreacionals com les platges o rius, o fins i tot, afectar masses d'aigua potable.

En el cas dels països més desenvolupats, on els tractaments de depuració de les aigües residuals eliminen la majoria de patògens bacterians, els virus representen els principals causants de brots relacionats amb la contaminació de l'aigua i aliments dels (Painter et al., 2013). Tradicionalment, i així es contempla en la legislació Europea vigent (Directiva 98/83/CE de l'aigua per al consum humà, Directiva 76/160/EEC de l'aigua de bany i la Directiva 91/271/CEE d'aigües residuals), es fan servir com a indicadors de la qualitat microbiològica de l'aigua *E. coli* i els enterococs intestinals, també anomenats *"Fecal Indicator Bacteria"* (FIB). Així doncs, amb la necessitat de complir la normativa vigent, els tractaments aplicats a les estacions depuradores d'aigua residual (EDAR) estan enfocats a reduir la càrrega microbiològica dels seus efluents, a partir de la determinació dels FIB com a criteri de qualitat microbiana. Altres microorganismes més persistents i resistents als tractaments utilitzats i al medi ambient, com per exemple els virus, no són inclosos en la normativa. Malgrat que la monitorització dels FIB permet detectar la presència de contaminació fecal bacteriana a l'aigua o els aliments, la seva absència no indica l'absència d'altres microorganismes patògens com virus o paràsits (Gerba et al., 1979; Jiang, 2006; Marzouk, 1980; Pusch et al., 2005; Savichtcheva and Okabe, 2006). Aquest fet ha suposat que aigües o aliments, teòricament salubres a través de la monitorització dels FIB, s'hagin vist implicats en brots (Dowell et al., 1995; Fleisher et al., 1996). Això és degut a que els bacteris són més sensibles als tractaments d'inactivació aplicats a les EDAR i a la radiació solar en comparació amb virus i protozous, disminuint la seva resistència a l'ambient (Enriquez, 1995; Madoni, 2011).Els principals virus transmesos per aigua i aliments tenen una ruta de transmissió fecal-oral i són: norovirus (NoV), rotavirus (RoV), virus de l' hepatitis A (VHA) i E

(VHE), adenovirus (AdV), astrovirus (AstV), sapovirus (SaV), poliomavirus (PyV) i enterovirus (EV). La seva concentració en períodes d'infecció aguda en femta oscil·la entre 10^8 i 10^{11} partícules víriques per gram (Lee et al., 2007). La monitorització de tots aquests patògens resulta inviable fent necessari l'ús d'indicadors que presentin una millor correlació amb patògens vírics. Per donar resposta a aquest problema s'han proposat mètodes basats en la detecció mol·lecular dels fags de *Bacteroides*, bacteriòfags d'RNA f-específics, colifags somàtics i fags d'*Enterococcus* spp. (Ahmed et al., 2008; Bonilla et al., 2010; Kirs and Smith, 2007; Ogorzaly et al., 2009). També s'ha suggerit l'ús de bacteris esporulats, molt més resistents al medi, com ara *C. perfringens* (Bisson and Cabelli, 1980). Virus humans com els adenovirus (Pina et al., 1998b; Puig et al., 1994), enterovirus (Noble et al., 2003; Noble and Fuhrman, 2001) i poliomavirus (Bofill-Mas et al., 2000) s'han utilitzat també com a indicadors de contaminació fecal. En aquest últim cas, els virus al ser específics d'hoste, permetrien traçar l'origen humà o animal de la contaminació. Els adenovirus humans (HAdV) i poliomavirus Humans (HPyV) compleixen moltes de les característiques exigibles a un bon indicador:

- Alta estabilitat sota condicions d'estrès ambiental com seria la radiació ultraviolada (UV), temperatura, concentració de clor, variació del pH i resistència als tractaments convencionals als que es sotmet l'aigua residual (Carter, 2005).

- Presència en aigües residuals durant tot l'any sense presentar estacionalitat i en concentracions elevades (Bofill-Mas et al., 2006). Els individus infectats per HAdV excreten el virus durant anys o mesos en femta sense produir símptomes clínics (Adrian et al., 1988; Pereira, 1972). El JCPyV s'excreta intermitentment en l'orina.

- Al ser específics d'espècie, els virus permeten identificar o suggerir d'on prové la contaminació fecal detectada, permetent traçar l'origen de la contaminació (Microbial Source Tracking, MST). En l'actualitat s'han descrit mètodes de quantificació d'adenovirus humans i porcins així com també de poliomavirus

humans, ovins i bovins (Bofill-Mas et al., 2006; Hernroth et al., 2002; Hundesa et al., 2010, 2009; Rusiñol et al., 2013).

1.4. Aigua residual: tractament i reutilització

Les aigües residuals urbanes són definides com masses d'aigua afectades per accions antropogèniques que n'han mermat la qualitat. Generalment aquestes aigües residuals són conduïdes, a través del sistema de clavegueram, a les EDAR on seran tractades. Els tractaments de depuració varien en funció de diversos factors: demogràfics, geogràfics, econòmics o de volum a tractar (Oller et al., 2011).

El fet que tots els individus d'una població contribueixin en la composició d'aquesta matriu a partir de l'excreció de femtes, orina o la descamació de la pròpia pell, fa que l'aigua residual sigui una matriu interessant per a realitzar estudis epidemiològics (Hellmér et al., 2014; La Rosa et al., 2013a).

1.4.1. Tractaments aplicats a les EDAR

Un cop a les EDAR, les aigües residuals són sotmeses a una sèrie de tractaments físics, químics i biològics que busquen eliminar els contaminants presents en l'aigua. Així doncs, l'objectiu dels tractaments és produir efluents més nets que puguin ser abocats a l'ambient, i un residu sòlid o fang (biosòlid o llot) que contindrà la majoria dels contaminants i nutrients. Els fangs de depuradora, un cop processats, poden ser utilitzats posteriorment com a adobs d'ús agrícola.

L'aigua residual crua constitueix una font d'entrada important de patògens al mediambient. Entre aquests patògens trobem bacteris, protozous, helmints i virus. Els virus presenten una alta estabilitat al medi ambient fet que facilita la seva transmissió (Carratalà et al., 2013; Enriquez, 1995; Rzeżutka and Cook, 2004).

A l'Estat Espanyol els paràmetres microbiològics i químics que ha de complir l'efluent d'una EDAR es recullen en el Reial Decret 1315/1992. Malgrat els

tractaments aplicats, existeixen nombrosos estudis que demostren la presència de virus als efluents de depuradora, suposant un risc per a la transmissió de malalties (Bofill-Mas et al., 2006; Da Silva et al., 2007; Pusch et al., 2005; Rodriguez-Manzano et al., 2012). Les etapes de les que consta el tractament d'aigües residuals es recull a la **Taula 1**.

Taula 1. Resum dels possibles tractaments aplicables per a depurar aigües residuals a les EDAR. (MO: Matèria Orgànica, DBO: Demanda Biològica d'Oxigen)

Fase	Objectiu principal	Mecanismes	Funció
Pretractament	Permet la separació de materials flotants de dimensions importants i l'eliminació de greixos	Desbast	Eliminació de grans objectes
		Desarenament	Eliminació de sorra
		Desgreix	Eliminació sòlids no miscibles
Primari	Sedimentació de materials flotants presents en la fracció líquida en un decantador. Reduir una tercera part la DBO	Sedimentació	Eliminació matèria en suspensió per acció gravetat
		Floculació	Formació de flòculs per acció de floculants
		Coagulació	Addicció de clorur fèrric o d'alumini que genera coàguls que ajuden a sedimentar
Secundari	Tractament biològic utilitzant bacteris que està enfocat a reduir la DBO de l'aigua un 85%. En moltes depuradores és l'etapa final del procés de depuració, generant efluents de qualitat acceptable	Llots actius	Bactèries degraden la matèria orgànica dissolta en un bioreactor
		Reactors Biològics MBR	Remouen tots els sòlids en suspensió
		Llacunatge	Eliminació per processos naturals (T, llum, biota) de patògens i matèria orgànica
		Filtres percoladors	Els microorganismes presents al filtre degraden la MO que s'hi diposita de l'aigua residual
		Biodiscs	Eliminen excés de matèria orgànica
		Sedimentació secundaria	Permet clarificar de matèria orgànica en suspensió
Terciari	Tractaments químics i físics avançats que tenen com a objectiu la millora significativa de la qualitat de l'aigua. Són tractaments necessaris per obtenir aigua regenerada o reutilitzada	Cloració	Desinfecció mitjançant hipoclorit
		Tractament amb UV	Inactivació de microorganismes
		Ozonització	Ús d'ozó per a desinfectar
		Llacunatge	Eliminació per processos naturals (T, llum, biota) de patògens i matèria orgànica
		Fotocatàlisi	Producció de radicals hidroxil que generen processos redox desinfectant
		Filtració amb llits de sorra	Eliminar patògens, olors i sabors
		Ultrafiltració	Eliminar molècules dissoltes de mida petita i col·loides
		Osmosi inversa	Eliminar ions inorgànics
		Precipitació química	Addicció de compostos químics que fan precipitar el fòsfor o reduir terbolesa
		Electrodiàlisi	Extracció selectiva de determinats ions

1.4.2. Aigües regenerades o reciclades

Les aigües regenerades o reciclades són aigües d'origen residual que han estat sotmeses a tractaments addicionals per millorar-ne la qualitat permetent la seva reutilització per a usos que no requereixen l'ús d'aigua potable. Així doncs, aquests tractaments permeten obtenir un nou recurs a partir d'un residu. Des de l'any 2007 existeix una normativa específica, el RD 1620/2007, que estableix el marc jurídic per a la reutilització d'aigües regenerades dins l'estat Espanyol. Aquesta norma ha representat un important avenç per estandarditzar les pràctiques de reutilització de l'aigüa sota 14 usos diferents agrupats segons 5 àrees principals: usos urbans, industrials, reacreacionals, ambientals i agrícoles. A Catalunya, l'aigua efluent tractada a les EDAR pot abocar-se al riu des d'on es captarà diluïda per a altres usos (reutilització indirecta o no planificada) o bé pot rebre un tractament terciari a les EDAR, sovint amb cloració o llum ultraviolada (reutilització directa o planificada), que permetrà millorar substancialment la seva qualitat microbiològica, i per tant, ampliar-ne el rang d'usos. Aquests usos venen determinats per criteris de qualitat, que a nivell microbiològic s'avaluen a través de la presència i quantificació d' *E. coli*. Per exemple, per a irrigar vegetals de fulla verda, on l'aigua regenerada entra en contacte amb la part comestible de la pròpia planta consumida en cru, calen aigües regenerades amb una concentració menor a 100 UFC/100 ml d'*E.coli*. Si l'aigua es fa servir per a regar arbres fruiters, el criteri microbiològic és més laxe, permetent fins a 10.000 UFC/100 ml d'*E.coli.*

A Catalunya, la demanda total d'aigua és de 3.123 hm^3 /any (ACA). Les projeccions estimen que de cara a l'any 2025 aquest consum incrementarà en uns 160 hm^3 addicionals (Guiu, 2000). Del total d'aigua consumida al nostre territori, el 72,6% (2.267 hm^3/any) es destina a usos ramaders i agrícoles mentre que el consum domèstic i industrial representa un 27,4% (856 hm^3/any). En les zones en què, per raons geogràfiques com pluviomètriques les aigües disponibles

o utilitzables són un factor limitant, l'aigua reutilitzada pot ser una alternativa. La producció i ús d'aigües regenerades permet un millor aprofitament dels recursos hídrics disponibles en contextos de baixa disponibilitat d'aigua. A la vegada, inclouen un component nutricional interessant que pot afavorir la producció agrícola (Dare, 2015; Wheaton et al., 2001). La producció d'aigua residual al nostre país ha anat en augment fins l'any 2008 (54hm^3/any) quan va assolir el seu màxim. L'any 2008, arran de la important sequera, la producció d'aigua regenerada va arribar al seu màxim històric. El Programa de Reutilització d'Aigua a Catalunya (PRAC) va marcar el 2015 com a objectiu, per arribar a regenerar el 22% de l'aigua residual produïda (ACA, 2009).

En la **Figura 3** pot observar-se la producció i usos d'aigua regenerada durant el període 2005-2015.

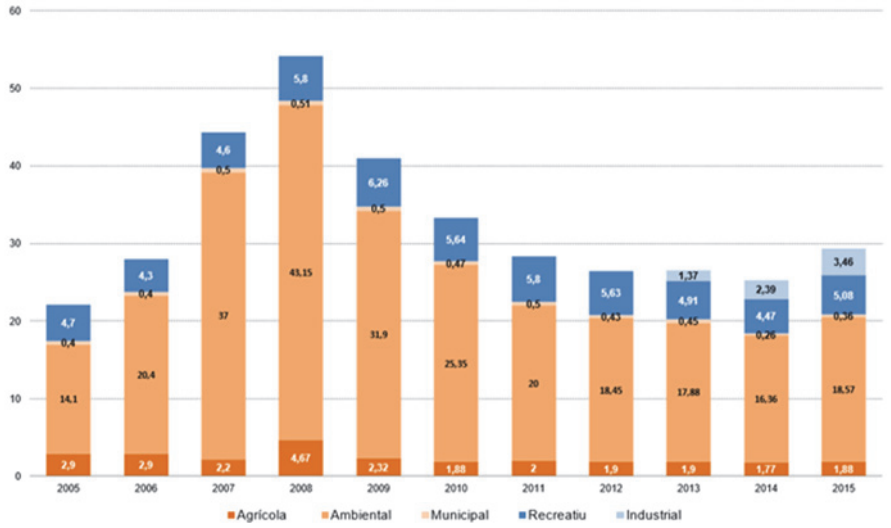


Figura 3. Producció d'aigua regenerada a Catalunya en els darrers anys. Font: dades oficials de l'ACA.

Segons dades de l'Agència Catalana de l'Aigua (ACA) del 2015, el consum d'aigua regenerada a Catalunya és d'aproximadament 30 hectòmetres cúbics. Aquesta xifra representa el 27% del cabal anual d'aigua tractat a les EDAR. Dels 30 hm^3

produits, la majoria ha estat destinat a usos mediambientals (63%) sent una petita part la destinada a usos agrícoles (2%) **(Figura 4)**.

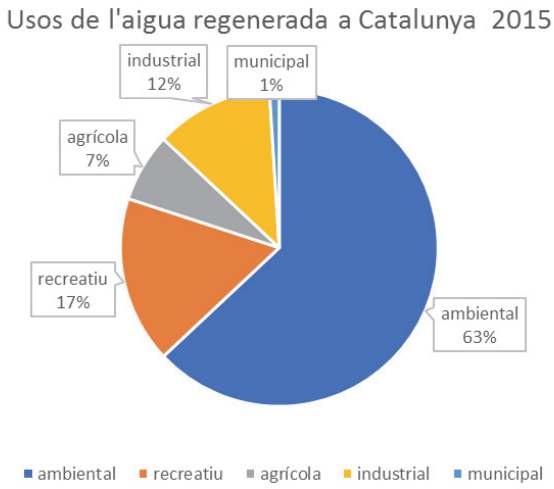


Figura 4. Usos de l'aigua regenerada durant l'any 2015. Font: dades oficials de l'ACA.

La vigent normativa d'aigües regenerades no estableix cap criteri de qualitat que faci referència als virus, malgrat que la manca de correlació entre els FIB i la presència/absència de virus és ampliament demostrada. Això suposa exposar la població a un risc no ben caracteritzat; determinats usos, com ara el reg de cultius que es consumeixen crus amb aigües regenerades, podria suposar un risc per la salut.

1.5. Els aliments com a font de transmissió de virus

El paper dels aliments com a vehícles de transmissió de malalties infeccioses es coneix des de l'antiguitat. Des d'un punt de vista únicament microbiològic, els aliments poden ser vehicle d'infeccions causades per protozous, bacteris, virus i prions. En el cas dels virus, les principals malalties associades al consum d'aliments contaminats són les gastroenteritis i les hepatitis, encara que altres malalties com les meningitis, miocarditis i altres alteracions neurològiques són possibles. Els virus entèrics ja són reconeguts com els principals agents de brots

alimentaris als EUA. Es calcula que dels 9.4 milions anuals de casos associats a brots alimentaris, el 58% d'ells estan relacionats amb norovirus, principal agent víric causant de gastroenteritis (Scallan et al., 2011). A nivell global s'estima que anualment 125 milions de brots alimentaris són causats per norovirus (Kirk et al., 2015).

Els aliments són susceptibles de ser contaminats a diferents nivells de la cadena de producció, començant per les fases d'irrigació i recol·lecció als camps, durant el seu pas per la industria alimentària, mentre són servits en locals de restauració o a la pròpia llar. Existeixen normatives que vetllen per la qualitat microbiològica dels aliments, tot i que aquestes no estableixen l'anàlisi de cap paràmetre víric. Així doncs, tot i les mesures de qualitat aplicades, els brots associats a virus són freqüents. Entre els principals aliments implicats en brots alimentaris d'origen víric trobem els mol·luscs bivalves. Importants brots epidèmics, afectant amb HAV a més de 292.301 persones, s'han reportat associats al consum d'ostres crues a Shangai, China (Halliday et al., 1991). Tot i que no tant importants, brots associats al consum de mol·luscs bivalves segueixen reportant-se contínuament (Polo et al., 2016; Woods et al., 2016). Altres aliments consumits crus també s'han associat a brots com ara els vegetals de fulla verda (Ethelberg et al., 2010; Müller et al., 2016) o fruites com les maduixes (Bernard et al., 2014; Mäde et al., 2013) o les magranes i gerds congelats (Collier et al., 2014; Sarvikivi et al., 2012). L'increment en la demanda de productes frescos, associats a una dieta sana, s'ha relacionat amb un increment dels brots vírics (Callejón et al., 2015). Kokkinos i col·laboradors (2012a) han estudiat la presència de HAdV, HAV, HEV, NoV GI i NoV GII en vegetals de fulla verda a tres països de la Unió Europea detectant-los en un 26,4% (70/265), un 0% (0/149), un 3,2% (4/125), un 1,3% (2/149), 0,8% (1/126), respectivament. La baixa prevalença de NoV detectats en aquest estudi contrasta amb l'exhaustiu estudi de Baert i colaboradors (2011) on es van detectar genomes de NoV entre el 28-50% dels vegetals estudiats. Finalment, una gran part dels brots alimentaris es donen en restaurants, altres establiments

o indústries on es manipulen aliments o s'ofereixen productes elaborats, els quals s'han manipulat de forma incorrecta. Els manipuladors d'aliments, ja sigui malalts o sense símptomes, poden contaminar accidentalment els aliments i causar brots (Barrabeig et al., 2010; Boxman et al., 2009; Nicolay et al., 2011).

Un dels punts crítics per evitar l'entrada de virus entèrics a la cadena alimentària és la utilització d'aigua de reg adequada. Recentment, un estudi publicat per Maunula i col·laboradors (2013) assenyalava que un 9,5% de les mostres d'aigua utilitzada per regar gerds contenia HAdV, demostrant la presència de contaminació fecal humana, i per tant, que la qualitat de l'aigua de reg no era la desitjable. En una línia similar, l'estudi de Kokkinos i col·laboradors va detectar la presencia de HAdV en el 27,9% de les mostres d'aigua de reg usada per a produir vegetals de fulla verda. Aquests exemples demostren que les aigües de reg suposen una important entrada dels virus a la cadena alimentària i que la principal mesura que podem prendre per evitar-los és utilitzar unes bones pràctiques agrícoles i de producció. Altres punts d'entrada dels virus a la cadena alimentària són els fertilitzants d'orígen animal o humà (Rodríguez-Lázaro et al., 2012). Recentment, s'ha especulat que la presència de virus en els aliments no només quedi limitada a la superfície, sinó que determinats virus patògens, com els NoV o els EV puguin ser internalitzats pel vegetal (Carducci et al., 2015; DiCaprio et al., 2015).

Actualment és el Reglament (CE) nº 2073/2005 i les seves posteriors modificacions (Reglamento (UE) 365/2010) els que fan referència als criteris microbiològics aplicables als productes alimentaris. En aquesta legislació no apareix cap menció específica a virus malgrat el perill que suposen. Aquest fet s'explica en part per la manca de mètodes sensibles, específics i estandarditzats per a concentrar i detectar virus en aliments. Recentment s'ha publicat la ISO 15216-1:2017 que estableix els mètodes per a concentrar i quantificar els virus de la HAV i els NoV en matrius alimentàries (https://www.iso.org/standard/65681.html). A l'haver solucionat la limitació

tècnica és d'esperar que la determinació de virus en algunes matrius alimentàries s'incloguin en futures revisions de la normativa.

1.6. Principals virus detectats en aigua i aliments

Les principals famílies víriques, que contenen patògens detectats per tècniques de seqüenciació massiva en les mostres ambientals estudiades en la present tesi doctoral, es detallen a continuació. Algunes d'aquestes famílies víriques han estat detectades també en mostres clíniques (**Figura 5**. Treball no inclòs a la tesi doctoral).

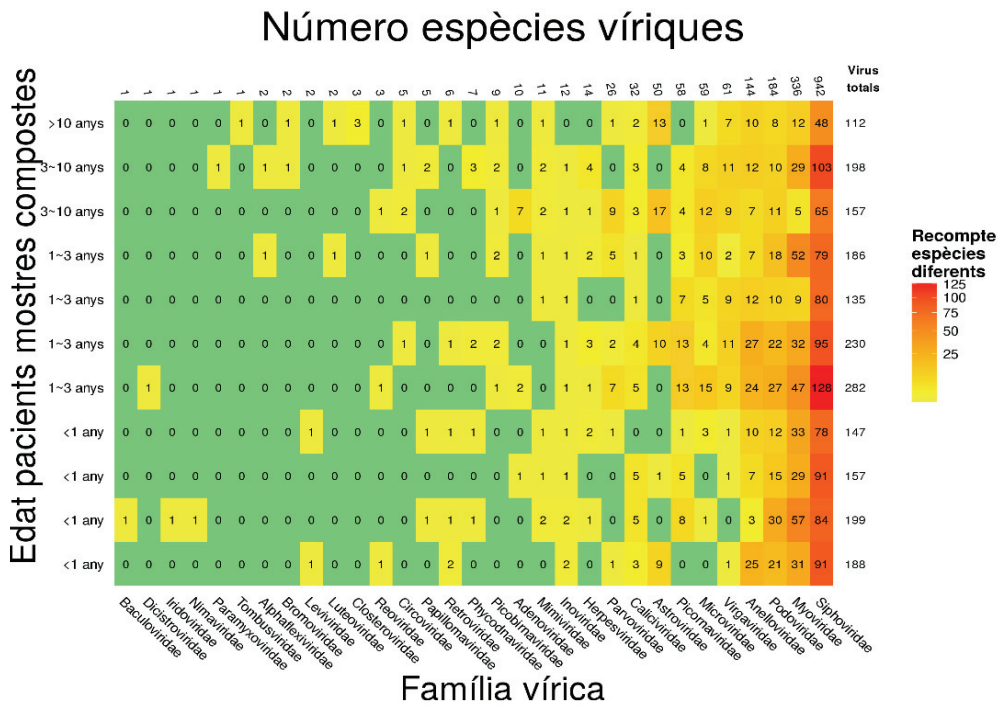


Figura 5. Famílies víriques detectades en mostres de pacients amb gastroenteritis agudes sense agent etiològic conegut

1.6.1. Adenoviridae

Els Adenovirus humans (HAdV) són virus que pertanyen a la família *Adenoviridae.* Els HAdV són virus icosaèdrics sense envolta amb un genoma lineal de doble cadena d'ADN d'unes 30-38 kb i un diàmetre de 60-90nm.

Els HAdV es troben classificats dintre del gènere *Mastadenoviridae* que infecta a mamífers. Altres gèneres dins de la família *Adenoviridae* són: *Atadenoviridae* (mamífers), *Aviadenovirus* (aus), *Ichtadenovirus* (peixos) i *Siadenovirus* (amfibis). Actualment existeixen més de 80 genotipus de HAdV diferents (http://hadvwg.gmu.edu/) basats en estudis serològics, seqüenciació completa del seu genoma i anàlisi filogenètiques. S'agrupen en 7 espècies diferents (A-G), sent els adenovirus de l'espècie D el grup més nombrós (Robinson et al., 2013). Els HAdV són capaços de provocar una gran diversitat de patologies com ara gastroenteritis (HAdV F), conjuntivitis (HAdV B,D i E), infeccions respiratòries (HAdV B , C, D i E), cistitis (HAdV B), meningitis (HAdV A, B i D) o hepatitis. Tradicionalment, els adenovirus també s'han classificat segons la capacitat de generar tumors en rosegadors (Figura 6) (Ginsberg, 1984).

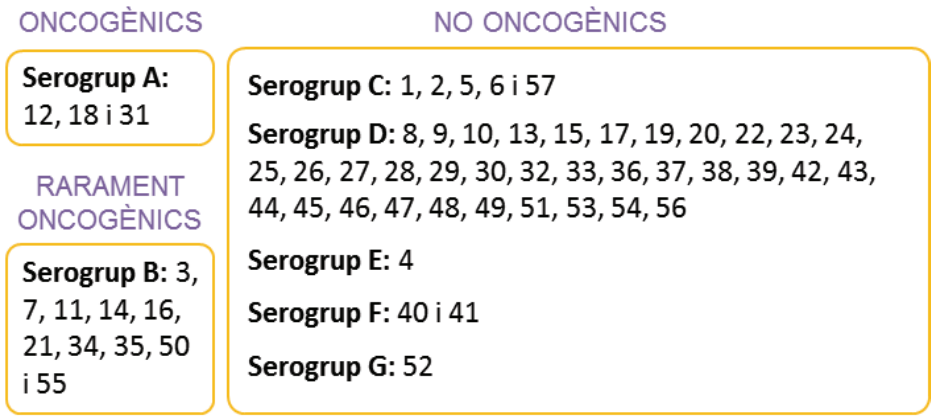


Figura 6. Serogrups i serotips d'adenovirus humans agrupats segons la capacitat de generar tumors en rosegadors. Figura adaptada a partir de Jones *et al*. (2007), Robinson *et al*.(2011) i Walsh *et al*. (2011)

Les infeccions per HAdV es troben molt esteses en la població i es produeixen de forma primerenca durant la infància. Estudis serològics en nens de diferents indrets geogràfics demostren la presència d'anticossos contra diferents genotipus d'HAdV en un 40-90% dels individus (D'Ambrosio et al., 1982; Chirmule et al., 1999; Thorner et al., 2006; Barouch et al., 2011). Els adenovirus produeixen

en molts casos infeccions subclíniques persistents que poden durar mesos o, fins i tot, anys. Aquest fet, juntament a la seva resistència als tractaments de desinfecció aplicats a les EDAR, a la seva detecció durant tots els mesos del any i a la seva prevalença en una gran diversitat d'àrees geogràfiques fa dels HAdV una eina útil per a traçar la contaminació fecal humana (Bofill-Mas et al., 2010, 2006; Dong et al., 2010; Hamza et al., 2009; Pina et al., 1998c).

1.6.2. Polyomaviridae

El primer poliomavirus (PyV) va ser descobert l'any 1953 en ratolins per Ludwing Gross. El seu descobriment es va produir a l'inocular homogeneïtzats de tumors de ratolins amb leucèmia a ratolins sans observant-se la formació de múltiples tumors (polioma). Els poliomavirus són virus sense embolcall lipídic amb un diàmetre de 50-60 nm. El seu genoma és ADN circular de doble cadena de 5,3 kb empaquetat amb 4 histones cel·lulars (H2A, H2B, H3 i H4) formant una estructura anomenada "microcromosoma". Ambdues cadenes de ADN són codificants i poden distingir-se tres regions diferenciades. La regió *early* o primerenca codifica per gens expressats abans de la replicació del virus (els antigens large T (LT) and small T (ST)); una regió tardana que codifica per proteïnes que s'expressen un cop s'ha replicat el genoma víric com les de la nucleocàpside (VP1, VP2, VP3 i LP1); i una regió reguladora que conté l'origen de replicació del virus. Els poliomavirus es classificaven en tres gèneres (*Avi-*, *wuki-* i *Orthopolyomavirus*) en funció de la similitud al llarg de tot el seu genoma. Un genoma amb una similaritat ≥19% es consideraba una nova espècie (Johne et al., 2011). El fet que alguns poliomavirus siguin recombinants ha fet que es proposi un nou sistema de classificació basat en divergències sobre la seqüència codificant de l'antigen Large T (LTAg), que ha de ser ≥15% (Calvignac-Spencer et al., 2016). Fins al dia d'avui, existeixen més de 1200 genomes complerts dins la família *Polyomaviridae* incloent més de 100 espècies diferents, la majoria de les quals s'han descobert en els últims 15 anys, sent 13 els poliomavirus que infecten l'ésser humà (Calvignac-Spencer et al., 2016). Els PyV produeixen infeccions durant la infància

i persisteixen durant mesos i fins i tot anys, excretant-se de forma perllongada en orina. Són virus molt resistents a les condicions de dessecació i molt resistents a l'ambient, detectant-se en totes les àrees geogràfiques durant tot l'any (Bofill-Mas et al., 2006). A la **Taula 2** apareixen tots els poliomavirus que infecten humans descrits fins ara.

Taula 2. Poliomavirus humans descoberts fins a l'actualitat. Adaptat a partir de Mishra i col·laboradors (2014)

Virus	Descobert en	Malaltia associada	Any	Referència
JCPyV	Cervell	Leucoencefalopatia multifocal progressiva (PML)	1971	Padgett et al., 1971
BKPyV	Ronyó		1971	Gardner et al., 1971
KI	Tracte respiratori		2007	Allander et al., 2007
WU	Tracte respiratori		2007	Gaynor et al., 2007
MCPyV	Pell	Carcinoma de cèl·lules de merkel	2008	Feng et al., 2008
HPyV-6	Pell		2010	Schowalter et al., 2010
HPyV-7	Pell		2010	Schowalter et al., 2010
HPyV-8	Pell	Tricodisplàsia espinulosa	2010	van der Meijden et al., 2010
HPyV-9	Sang Ronyó		2011	Scuda et al., 2011
HPyV-10, MLXPyV MWPyV	Femta	WHIM	2012	Buck et al., 2012; Siebrasse et al., 2012
HPyV-11	Femta		2013	Lim et al., 2013
HPyV-12	Teixit hepàtic		2013	Korup et al., 2013
HPyV-13 o NJSPyV	Endoteli		2014	Mishra et al., 2014

1.6.3. Caliciviridae

Els membres de la família *Caliciviridae* són virus de mida petita de 27-40 nm, sense envolta, amb un genoma de cadena ARN de polaritat positiva i simetria icosaèdrica. La família *Caliciviridae* conté 5 gèneres diferents*: Norovirus, Sapovirus, Lagovirus, Nebovirus i Vesivirus* (Freed and Martin, 2013). La inclusió dels gèneres *Bavovavirus, Nocavirus, Recovirus, Valovirus i Secalivirus* descrits recentment està en avaluació (Farkas et al., 2008; L'Homme et al., 2009; Ng et al., 2012; Wolf et al., 2012, 2011).

Dintre dels gèneres *Norovirus* i *Sapovirus* trobem importants agents causals de gastroenteritis que infecten humans. La **Figura 7** representa l'organització dels ORF en els membres de la família *Caliciviridae*.

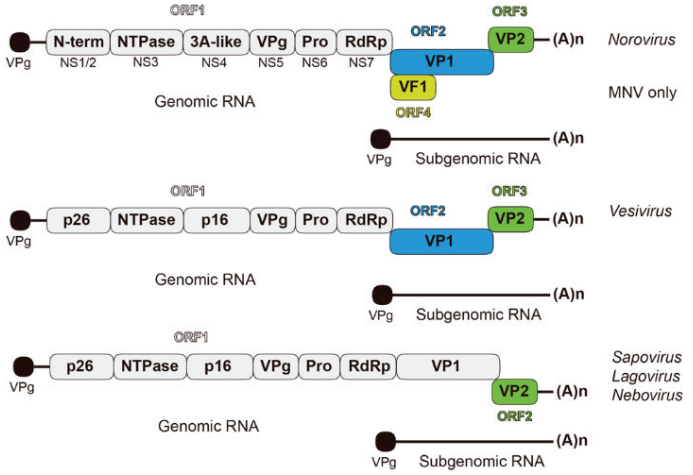


Figura 7. Representació esquemàtica dels ORF dels diferents membres de la família *Caliciviridae* extret de Royall and Locker (2016)

1.6.3.1. Gènere *Norovirus*

Els norovirus humans (HNoV), prèviament coneguts com a virus de Norwalk, van ser identificats a partir de mostres de femta associades a un brot a la ciutat americana de Norwalk, considerant-se el primer agent etiològic víric causant de gastroenteritis (Kapikian et al., 1972). Els norovirus humans tenen un genoma

lineal ARN de polaritat positiva d'unes 7,6 kbs. El seu genoma es troba covalentment lligat a una proteïna (VPg) en l'extrem 5' i poliadenilat a l'extrem 3' (Lambden et al., 1993). Conté tres pautes de lectura oberta que codifiquen per 8 proteïnes en total. ORF2 i ORF3 codifiquen per les proteínes estructurals del virus (VP1 i VP2). L'ORF1 codifica per una poliproteïna que es processa donant lloc a 6 proteïnes no estructurals, entre elles la ARN-polimerasa ARN-depenent del virus (Thorne and Goodfellow, 2014). La seqüència completa de la proteïna VP1 permet la classificació dels norovirus en 6 genogrups diferents (GI-GVI) i més de 40 genotipus diferents (Zheng et al., 2006). La majoria de brots de gastroenteritis en humans estan provocats per els genogrups II en concret per la soca GII.4 (Vega et al., 2011; Verhoef et al., 2015) i el genotipus II.17 que està desplaçant progressivament al GII.4 com a soca circulant (de Graaf et al., 2015; Koo et al., 2017). Recentment s'ha descrit un genotipus recombinant emergent GII.P16-GII.2 (Niendorf et al., 2017). Esporàdicament hi ha brots associats al genogrup I i més rarament al genogrup IV. Segons estimacions del Centers for Disease Control (CDC) els norovirus són responsables del 60% de gastroenteritis agudes amb causa coneguda. Es calcula que els NoV són responsables de la mort d'uns 200.000 nens menors de 5 anys en països en vies de desenvolupament, sent la segona causa de gastroenteritis en nens després dels rotavirus (Patel et al., 2008). Estudis de revisió han demostrat que el NoV GII.4 està més associat a transmissió persona-persona, especialment en ambients hospitalaris o residències, mentre que els genotipus GI.7 i GII.12 s'associa a transmissió d'origen alimentari (Vega et al., 2014). A diferència d'altres virus, la immunitat adquirida per infeccions prèvies de norovirus no es perllonguen en el temps (Karst et al., 2003). S'han observat diferents susceptibilitats en front a l'infecció de NoV associades a l'existència de polimorfismes en els gens ABO, FUT2 i FUT3 que codifiquen per glucosiltransferases. Aquests gens són els responsables d'afegir molècules de fucosa que intervenen en el procés d'unió del virus al receptor condicionant la susceptibilitat dels individus en front a la infecció

(Ruvoën-Clouet et al., 2013). Els NoV estan en constant evolució, apareixent contínuament noves soques degut a mutacions puntuals (deriva genètica) o a fenòmens de recombinació durant una co-infecció (Bull et al., 2007). Els fenòmens de recombinació tenen un paper important en l'aparició de noves soques antigènicament diferents capaces d'evadir la resposta immune (White, 2014). La zona d'unió entre l'ORF1 i ORF2 del genoma de NoV s'ha identificat com a punt freqüent de recombinació (Arana et al., 2014; Fumian et al., 2016).

L'aparició de brots de gastroenteritis associats a norovirus ha estat àmpliament descrita a la literatura. Recentment hi ha hagut a Catalunya un brot que ha afectat a més de 4.000 persones i que s'ha associat al consum d'aigua embotellada contaminada. Altres brots s'han associat al consum d'aigua municipal potable (Kaplan et al., 1982; Maunula et al., 2005; Riera-Montes et al., 2011), al consum de gel contaminat (Cannon et al., 1991) o a l'exposició d'aigües recreacionals contaminades (ter Waarbeek et al., 2010). Els norovirus també han provocat diversos brots associats al consum d'aliments contaminats. Els principals aliments associats a brots de norovirus són vegetals de fulla verda (Ethelberg et al., 2010), bivalves (Baker et al., 2011; Lowther et al., 2012; Nenonen et al., 2008), fruits vermells (Bernard et al., 2014; Mäde et al., 2013) o productes de pastisseria (Guo et al., 2014).

1.6.3.2. Gènere *Sapovirus*

Els sapovirus (SaV) tenen un genoma ssRNA de polaritat positiva amb dos ORFs a diferència dels NoV. L'ORF1 codifica per proteïnes no estructurals i la proteïna major de la càpsida VP1 mentre que l'ORF2 codifica per la VP2. La proteïna VP1 és la regió amb major diversitat del genoma víric (Oka et al., 2012) i es fa servir per classificar els SaV en 5 genogrups (GI a GV) existint uns altres 8 proposats pendents d'aprovació (GVI a GXIV) (Scheuer et al., 2013). De tots els genogrups només 4 genogrups infecten a Humans (GI, GII, GIV i GV). Els SaV produeixen gastroenteritis similars a les provocades per NoV. De fet en algunes regions s'ha identificat la infecció per HSaV com la tercera causa més comú de

gastroenteritits després dels rotavirus i norovirus (Iritani et al., 2016). Malgrat tot, el nombre de brots relacionats amb SaV és menor al nombre de casos reportats per NoV (Torner et al., 2016). S'han descrit brots en creuers, hospitals, escoles i residències. També s'han descrit brots d'origen alimentari implicant mol·luscs bivalves (Hansman et al., 2008a; Iizuka et al., 2010). Un dels majors brots per SaV ha estat descrit al Japó al 2010 afectant a més de 600 persones que consumien menjar preparat d'estil japonès (Kobayashi et al., 2012). Els SaV circulen per l'ambient a altes concentracions (Haramoto et al., 2008; Sano et al., 2011).

1.6.4. Hepeviridae

Els virus de la família *Hepeviridae* són virus sense embolcall, amb càpsida icosaèdrica i un diàmetre de 32 a 34 nm. El seu genoma és una cadena senzilla de RNA amb polaritat positiva d'unes 7,2 kb (Ahmad et al., 2011). Aquesta família conté un important patògen humà com és el virus de l'Hepatitis E (VHE). El genoma del VHE conté tres ORF superposats. L'ORF1 codifica per les proteïnes no estructurals del virus, l'ORF2 codifica per les proteïnes de la càpsida mentre que l'ORF3 codifica per una proteïna que es creu implicada en la sortida dels virions de la cèl·lula infectada.

Actualment la família *Hepeviridae* conté dos gèneres: *Piscihepevirus* i *Orthohepevirus.* El gènere *Piscihepevirus* conté una única espècie que infecta truites (Batts et al., 2011). El gènere *Orthohepevirus* conté 4 espècies diferents que infecten diferents hostes com aus (Orthohepevirus B), carnívors (Orthoepevirus C) i ratpenats (Orthohepevirus D). L'espècie Orthohepevirus A conté 7 genotipus diferents que infecten un ampli rang d'hostes (**Figura 8**).

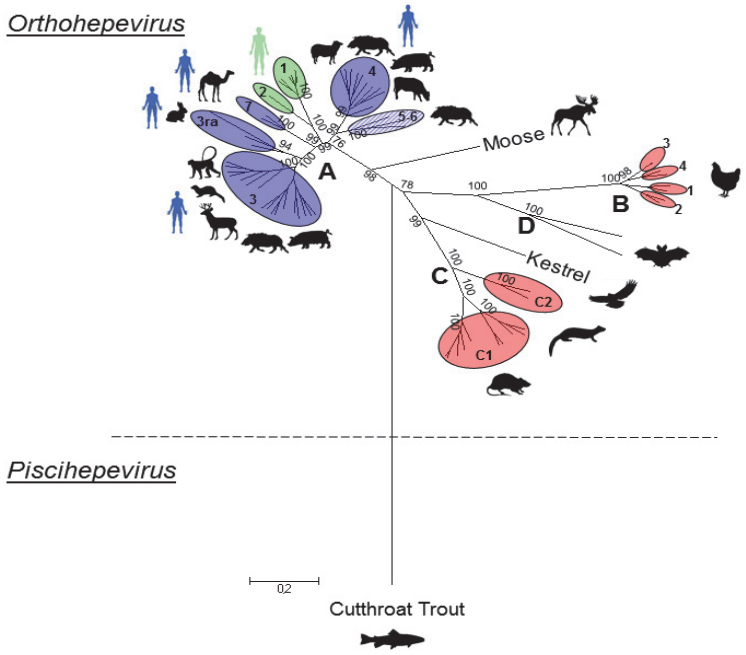


Figura 8. Arbre filogenètic amb els diferents gèneres i genotips dins la família *Hepeviridae* extreta de Doceul i col·laboradors (2016)

Els humans poden patir infeccions associades als genotipus 1, 2 (considerats principalment humans) i els genotipus 3, 4 i 7 (considerats zoonòtics) (Legrand-Abravanel et al., 2009). Altres genotipus infecten varietat d'animals com ara els senglars (VHE 3, 4, 5 i 6), els cèrvols (VHE-3), els camells (VHE-7), els porcs (VHE-3 i 4) i els conills (VHE-3) (Van der Poel et al., 2016).

Els diferents genotips que infecten humans presenten una distribució geogràfica diferenciada. Els genotips 1 i 2 són predominants en zones amb baix nivell de

sanejament de l'aigua. En canvi, els genotips 3 i 4, considerats zoonòtics, són predominants a la resta de zones (Kamar et al., 2012). La via de transmissió més freqüent dels genotips 3 i 4 als humans està associada al consum d'aliments contaminats i no hi ha descrits brots epidèmics importants (Lewis et al., 2010).

El VHE és considerat un patògen emergent que es transmet per la via fecal-oral, i per tant, la seva transmissió pot donar-se al consumir aigua o aliments contaminats. En la majoria dels casos el virus de la VHE causa infeccions autolimitades que tenen una durada d'unes 2 setmanes. La simptomatologia clínica pot incloure els següents signes, símptomes i síndromes: hepatomegàlia, febre, nàusees, vòmits, icterícia, colúria i anorèxia. Els símptomes i signes apareixen aproximadament unes 5 o 6 setmanes després de la infecció. Aquest decalatge de temps entre la infecció i l'aparició dels símptomes dificulta el traçat dels brots. Segons estimacions de la OMS es calcula que existeixen uns 20 milions d'infeccions anuals relacionades amb el consum d'aigua contaminada. Aproximadament uns 3,3 milions de casos són simptomàtics causant unes 56.600 defuncions (OMS., 2015). Els brots de VHE s'han associat sovint a situacions de crisi humanitària (Boccia et al., 2006; Guerrero-Latorre et al., 2011). Malgrat això, s'han descrit brots d'origen alimentari associats al consum de carn contaminada (Li et al., 2005; Yazaki et al., 2003). El virus també ha estat detectat en marisc suggerint una possible via de transmissió a través del seu consum (Gao et al., 2015; Grodzki et al., 2014).

1.6.5. Picornaviridae

Els virus de la família *Picornaviridae* són virus esfèrics d'uns 30 nm, sense embolcall, amb un genoma ARN de cadena senzilla amb polaritat positiva d'entre 7,0 - 8,8 kbs.

Aquesta família inclou importants patògens humans com el virus de l'hepatitis A, els enterovirusi el virus de la malaltia de mans, peus i boca o del refredat comú. A març del 2017 la família *Picornaviridae* inclou 35 gèneres diferents amb més de 80 espècies reconegudes pel grup de taxonomia de la ICTV (http://www.picornaviridae.com/). El gènere *Hepatovirus* inclou una única espècie: el virus de l'hepatitis A, un dels principals agents etiològics d'hepatitis transmesa a partir del consum d'aigua i aliments contaminats. Existeixen sis genotips de HAV amb distribució mundial, sent els genotips I, II i III els que principalment infecten humans (Vaughan et al., 2014). En zones del món on els nivells de sanejaments són baixos, els estudis de prevalença indiquen que aproximadament el 100% de la població infantil adquireix immunitat durant la primera dècada de vida. En societats amb sistemes de sanejament més desenvolupats s'ha reduït la seroprevalença en persones joves i, per tant, en regions on no està sistematitzada la vacunació, part de la població és susceptible d'infectar-se si entren en contacte amb el virus (Jacobsen and Wiersma, 2010). A la **Figura 9** pot observar-se l'organització genòmica del virus de la HAV com a model de picornavirus.

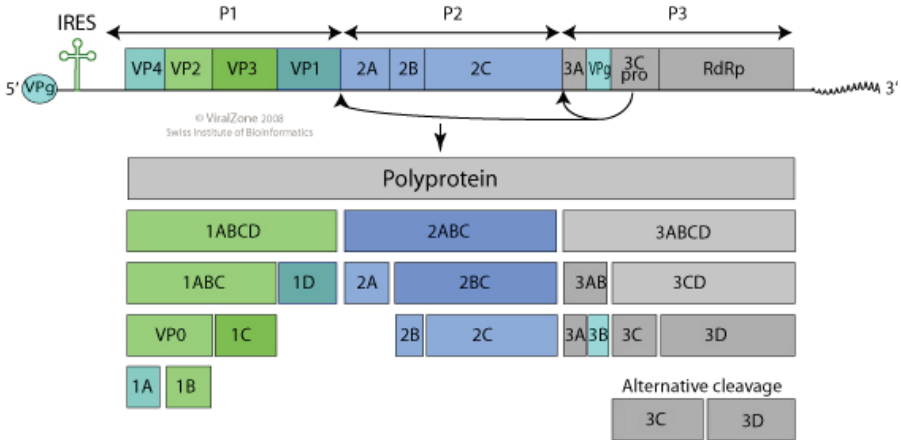


Figura 9. Organització genòmica del virus de l'hepatitis A com a exemple de picornavirus. Imatge obtinguda de Viral Zone: http://viralzone.expasy.org

El gènere *Cardiovirus* conté dues espècies conegudes com els virus de Theilor i el virus de la encefalomiocarditis. Els virus de la encefalomiocarditis són virus de rosegadors que poden infectar diversos hostes incloent l'home on s'ha associat a l'aparició d'un quadre frebril amb nàusees, mal de cap i dispnea (Oberste et al., 2009). L'espècie Theilorvirus inclouria els virus de Saffold que infecten a éssers humans malgrat que el seu paper com a virus patògens és controvertit (Li et al., 2017; Nielsen et al., 2012, 2013).

El gènere *Enterovirus* conté 10 espècies diferents incloent els Enterovirus humans A-D així com els Rhinovirus humans A-C. El gènere *Enterovirus* s'ha associat a l'aparició de malalties similars a la pòlio, malaltia de mans, peus i boca i a quadres gastrointestinals (Plyusnin et al., 2011). Recentment, hi ha hagut un increment en el nombre de brots associats a EV. Alguns d'aquests brots han estat associats a soques emergents recombinants (Holm-Hansen et al., 2016; Zhang et al., 2010). Són d'especial importància l'EV-71 i l'EVD-68. L'EV-71 conté 11 subgenogrups diferents, i en concret, el subgenogrup C4 s'ha associat a desordres neurològics importants amb una alta mortalitat (Fischer et al., 2014). L'any 2016 hi va haver un important brot d'encefalitis infantil principalment causada per EV-71 subgenogrup C1 a Catalunya (Casas-Alba et al., 2017). L'EV-D68 s'ha associat a patologia respiratòria greu i a paràlisi flàccida (Robinson et al., 2014). Els EV es transmeten principalment per contagi persona-persona però s'han descrit brots a la literatura lligats al consum d'aigua (Beller et al., 1997; Häfliger et al., 2000) o d'aliments contaminats (Le Guyader et al., 2008). Dintre del gènere *Enterovirus*, trobem els rinovirus humans, responsables del refredat comú, i el virus de la pòlio, patògen històricament important responsable de la poliomielitis. Aquesta malaltia, que afecta a nivell del sistema nerviós central destruint les neurones i produint paràlisi aguda flàccida o respiratoria, va ocasionar una greu epidèmia durant els segles XIX i XX. L'any 1988 l'Assemblea Mundial de la Salut va llançar l'iniciativa mundial d'erradicació de la pòlio (GPEI, del anglès *Global Polio Erradication Initiative*) que promovia programes de

vacunació i vigilància intensius. Gràcies a aquesta iniciativa l'incidència de poliomielitis s'ha reduït en més d'un 99% i únicament queden tres països on la malaltia és encara endèmica: Afganistan, Pakistan i Nigeria. (Zaidi et al., 2016).

El gènere *kobuvirus* conté patògens humans com els virus d'Aichi que s'han relacionat amb patologia gastrointestinal (Yamashita et al., 1991) però estudis recents senyalen que podrien tenir un paper en co-infeccions amb altres virus (Ambert-Balay et al., 2008; Räsänen et al., 2010). El virus d'Aichi s'ha detectat en un ampli ventall de mostres ambientals: aigua residual, aigua tractada i aigua superficial així com en mol·luscs bivalves (Hansman et al., 2008b; Kitajima et al., 2011; Lodder et al., 2013). Els parechovirus humans són patògens que es classifiquen dintre del gènere *Parchovirus* provocant patologia respiratòria i gastrointestinal (Chen et al., 2014). Els parechovirus també s'han relacionat amb afectacions del sistema nerviós central especialment en nens (Felsenstein et al., 2014). Recentment, l'aplicació de tècniques de seqüenciació massiva ha permès identificar multitud d'espècies noves que esperen ser reconegudes per part del grup de taxonomia de la família *Picornaviridae* a l'ICTV. Alguns d'aquests picornavirus s'han relacionat amb infeccions gastrointestinals en humans com és el cas dels gèneres *Cosavirus*, *Salivirus/Klassevirus* que han estat aïllats de femtes de pacients malalts (Holtz et al., 2009a, 2008).

L'utilització de tècniques de seqüenciació massiva al camp de la virologia ha permès el descobriment de noves espècies virals que podrien representar potencials nous gèneres dins de la família *Picornaviridae* (Boros et al., 2013; Holtz et al., 2009b; Lim et al., 2014; Ng et al., 2015; Phelps et al., 2014).

1.6.6. Astroviridae

Els astrovirus són virus amb càpsida icosaèdrica, sense embolcall i amb una mida de 28-41 nm. El seu genoma és RNA de cadena senzilla amb polaritat positiva d'unes 6,8 kbs (6,2-7,8 kbs) (Freed and Martin, 2013) contenint tres ORFs, coneguts com a ORF1a, ORF1b, ORF2. En l'extrem 5' trobem els ORF1a i ORF1b

que codifiquen per proteïnes no estructurals implicades en la replicació del virus mentre que a l'extrem 3' trobem l'ORF2 que codifica per les proteïnes estructurals del virus (Monroe et al., 1993; Willcocks and Carter, 1993). La família *Astroviridae* conté dos gèneres, el gènere *Mamastrovirus* (MAstV) que infecta a una gran diversitat de mamífers, entre ells els éssers humans, i el gènere *Avastrovirus* (AAstV) que infecta a aus (Figura 10). Inicialment la classificació dins del gènere depenia únicament de l'hoste infectat pel virus però, en els últims anys s'han descobert nous astrovirus que infecten un ampli rang d'hostes diferents invalidant el sistema de classificació anterior (De Benedictis et al., 2011). Això ha fet que per a classificar-los a més de l'hoste s'avaluin les diferències genètiques existents en la seqüència completa de la nucleocàpsida. Dintre del gènere MAstV trobem diversos genotips que infecten humans: MAsTV1 (inclou els clàssics serotips HAsTV1-8), MAsTV-6 (HAstV-MBL), MAstV-8 i MAsT-9 (HAstV-VA/HMO)(Bosch et al., 2014a).

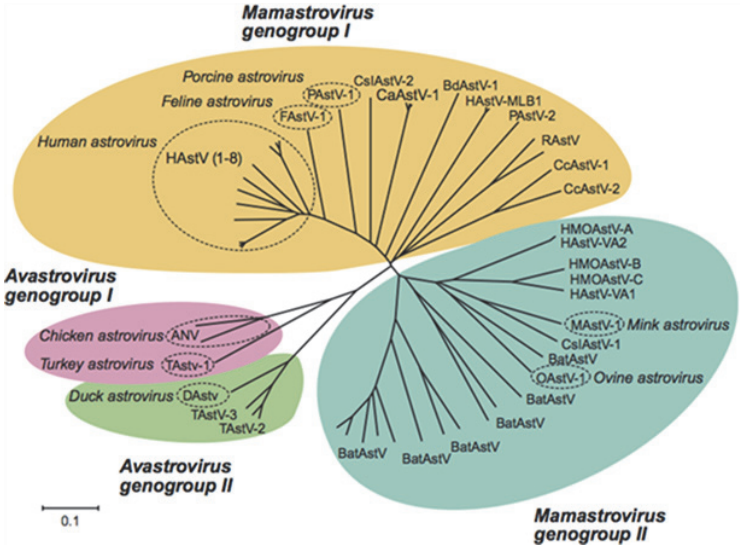


Figura 10. Arbre filogenètic de la família *Astroviridae* utilitzant la seqüència completa de la nucleocàpisda dels virus ("Family – Astroviridae," 2012)

Els astrovirus humans inclosos en el gènere MAstV-1 són responsables de gastroenteritis i s'estima que causen, a nivell mundial, entre el 2-9% de les gastroenteritis no bacterianes en nens (Bosch et al., 2014b). Els genotips

recentment descrits com els MAstV-6 s'han associat a diarrea (Finkbeiner et al., 2008) però també s'han aïllat en sang i nasofaringe de pacients amb infecció al tracte respiratori superior, suggerint un possible paper en altres patologies (Holtz et al., 2011; Wylie et al., 2012). Els MAstV-8 i 9 s'han associat a gastroenteritis (Finkbeiner et al., 2009a) i s'ha observat el seu tropisme al sistema nerviós central en nens inmunodeprimits que presentaven encefalitis (Quan et al., 2010). La transmissió fecal-oral dels astrovirus ha estat demostrada en estudis amb voluntaris (Kurtz et al., 1979). Es tracta de virus altament prevalents a l'ambient (Guimarães et al., 2008; Le Cann et al., 2004; Pusch et al., 2005), i són molt estables en aigua de beguda o aigües superficials (Bosch et al., 1997). Malgrat que s'han reportat brots d'astrovirus a través del consum d'aigua contaminada i aliments (Oishi et al., 1994; Pintó et al., 1996) la seva freqüència és menor comparada amb els brots associats a NoV. Com a exemple, la base de dades de vigilància del CDC dels Estats Units només ha reportat dos brots d'astrovirus des de 1998. Aquesta baixa taxa de brots podria ser deguda a la manca de monitoreig rutinari dels HAst en mostres clíniques i ambientals. A més a més els mètodes convencionals basats en PCR podrien fallar a l'hora de detectar HAstV degut a l'alta taxa de recombinació (Walter et al., 2001) i l'alta taxa de mutació (Babkin et al., 2012) pròpies d'aquesta família. L'aplicació de tècniques de seqüenciació massiva no basades en encebadors específics ha permès la detecció de multitud d'AstV recombinants que poden infectar a humans (Finkbeiner et al., 2009b; A. Kapoor et al., 2009). Futurs estudis per conèixer la prevalença d'astrovirus clàssics i recombinants són necessaris.

1.6.7. Reoviridae

Els rotavirus reben el seu nom del llatí "*rota*" que significa roda degut al seu aspecte al microscopi electrònic. La família *Reoviridae* conté 15 gèneres diferents dividits en dues sub-famílies en funció del seu aspecte morfològic: *Sedoreovirinae* (sense espícules) i *Spinareovirinae* (amb espícules). Dintre del primer grup trobem els gèneres *Rotavirus* (vertebrats), *Cardoreovirus* (crustacis),

Mimoreovirus (protistes marins), *Orbivirus* (mamífers, artròpodes i ocells), *Phytoreovirus* (plantes i insectes), i *Seadornavirus* (humans i insectes). Dins de la sub-familia *Spinareovirinae* trobem els gèneres *Aquareovirus* (peixos i moluscs), *Coltivirus* (mamífers i artròpodes), *Cypovirus* (insectes), *Dinovernavirus* (insectes), *Fijivirus* (plantes i insectes), *Idnoreovirus* (insectes), *Mycoreovirus* (fongs), *Orthoreovirus* (vertebrats), i *Oryzavirus* (plantes i insectes). Com es pot observar, aquesta família vírica és molt diversa i infecta un ampli rang d'hostes.

Els rotavirus humans (HRoV) són la principal causa de gastroenteritis severa amb una major mortalitat en infants menors de 5 anys. S'estima que unes 500.000 morts anuals són causades per RoV afectant especialment a països de baixa renda on l'accés a l'aigua és limitat (Tate et al., 2012). Els RoV és classifiquen serològicament en 7 grups (A, B, C, D, E, F, G) en funció de la resposta serològica envers la proteïna VP6.

El 90% de les infeccions de rotavirus estan causades per RoVA (Dennehy, 2000) malgrat que els grups B i C també poden infectar humans. Els HRoV són virus no embolcallats, amb una càpsida icosaèdrica i 70 nm de diàmetre. El seu genoma és de doble cadena de RNA segmentat en 11 fragments que codifiquen per 6 proteïnes estructurals (VP1-VP4, VP-6 i VP-7) i 6 proteïnes no estructurals (NSP1-NSP5/6) (**Figura 11**) (Freed and Martin, 2013).

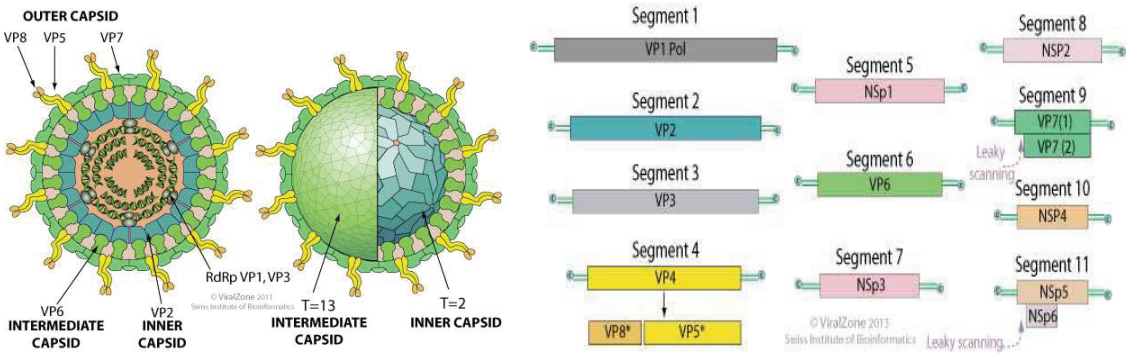


Figura 11. Presentació de la partícula vírica i ORFs dels rotavirus humans. Font: Viral Zone, http://viralzone.expasy.org

A dia d'avui existeixen 27 genotips G i 37 genotips P. La majoria de les infeccions humanes per Rotavirus s'associen a 5 genotips G (GI-G4 i G9) i 3 genotips P (P[8], P[4] i P[6]) (Matthijnssens et al., 2011; Santos and Hoshino, 2005). La transmissió dels rotavirus és per la via fecal-oral i es requereix una quantitat molt baixa de partícules víriques per a manifestar clínicament la malaltia (Ward et al., 1986). Al tractar-se d'un virus sense embolcall és altament estable en l'ambient i a més el fet que s'eliminin fins a 10^{12} partícules víriques per gram de femta facilita la seva transmissió a través de fomites contaminats (Bishop, 1996; Estes et al., 1979). Els rotavirus tenen una distribució geogràfica mundial. Prèviament a la introducció de la vacuna en els països desenvolupats, els RoV presentaven un marcat pic estacionari a l'hivern (Parashar et al., 1998). La seva detecció ambiental s'ha descrit tant en aigua residual crua (Barril et al., 2015; Prado et al., 2011; Zhou et al., 2016) com en aigua de riu (Calgua et al., 2013a; Prez et al., 2015). El nombre de brots alimentaris associats exclusivament a RoV són poc freqüents (CDC, 2000; Mayr et al., 2009). En la majoria de casos es tracta de co-infeccions amb altres patògens com ara els NoV (RÄSÄNEN et al., 2010). Existeixen diverses vacunes atenuades per a RoV, a l'Estat Espanyol estan disponibles Rotarix® (GSK) i Rotateq® (Merck). Rotarix® conté un virus atenuat G1P[8] i ha demostrat una

eficiència del 85-96% contra gastroenteritis moderades-severes causades per RoV serotips G1 i altres genotips (O'Ryan, 2007). Rotateq®, en canvi, és una vacuna pentavalent generada per reordenament de rotavirus boví i els serotips humans G1, G2, G3, G4 i P[8]. Rotateq® ha demostrat una eficiència del 74% (66,8-79,9) durant el primer any post-vacunació contra la gastroenteritis causada pels genotipus G1-G4 (Vesikari et al., 2006).

1.6.8. Papillomaviridae

La família *Papillomaviridae* (PV) està constituïda per virus circular de doble cadena ADN d'unes 7-8 kbs que infecten vertebrats. Es tracta de virus amb simetria icosaèdrica sense embolcall d'uns 50-60 nm. Només una de les dues cadenes de ADN es tradueix a proteïnes seguint el sentit horari i les seves pautes de lectura oberta es classifiquen en primerenques (E1-E7) o tardanes (L1, L2) en funció de la seva posició al genoma. Existeix una regió no codificant important per a la replicació del virus i la regulació d'expressió de proteïnes. Existeixen més de 29 gèneres de papil·lomavirus, essent 5 d'ells exclusivament humans i classificant-se en funció de la seqüència nucleotídica completa del gen L1 (Bernard et al., 2010). En els últims anys el nombre d'espècies diferents que pertanyen a la família *Papillomaviridae* ha incrementat considerablement. En concret, existeixen més de 200 genotips diferents que infecten l'ésser humà (**Figura 8**) (http://www.hpvcenter.se/html/refclones.html). Els genotips pertanyents a un mateix gènere tenen una similitud en L1 ≥ al 60% entre ells. S'anomena espècie a aquells HPV que comparteixen una identitat entre 60-70% en L1. Els genotips de HPV comparteixen una similitud entre 71-89% sobre la seqüencia completa de L1.

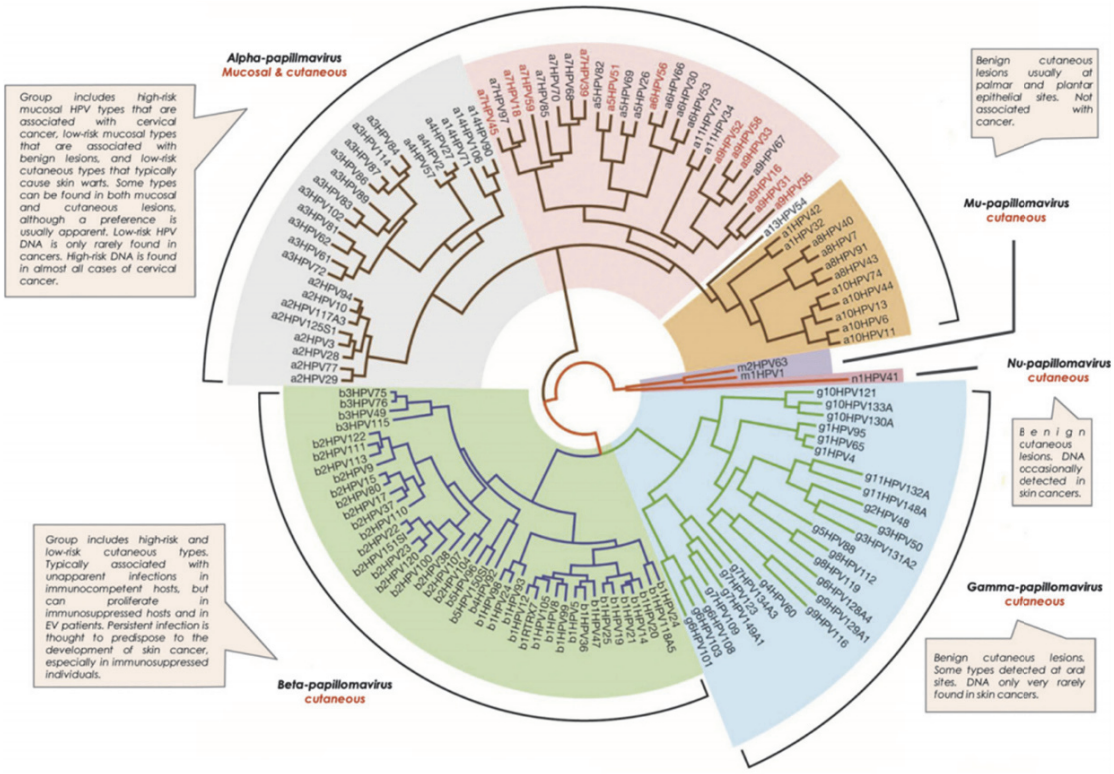


Figura 12. Arbre filogenètic de la classificació dels HPV en funció de la proteïna L1 (Doorbar et al., 2012)

Les proteïnes E6 i E7 estan estretament relacionades amb les propietats oncogèniques dels HPV al bloquejar les proteïnes de retinoblastoma i p53, induint l'entrada de la cèl·lula en fase S (Werness et al., 1990). Els papil·lomavirus humans alfa (HPV- α) infecten principalment la superfície de les mucoses orals i genitals així com també els genitals externs. Els altres gèneres de HPV β, γ, μ, i ν infecten les mucoses no genitals i la pell. Entre els 120 genotips diferents que pertanyen al gènere α, 30 infecten l'epiteli anogenital i són causa de malalties de transmissió sexual com ara els condilomes. Dintre del gènere α (HPVα) 12 genotips tenen elevat potencial oncogènic i es consideren d'alt risc (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, i 59) sent reconeguts com agents causants de càncer de cèrvix (CC), el segon càncer més freqüent en dones a nivell mundial (Muñoz et al., 2003). La gran majoria de casos de CC s'associen al genotips HPV16

(61% dels casos). Els genotips de HPV16 i 18 també s'han associat a altres càncers com el d'anus, vulva i penis. Recentment també s'ha associat la presència de HPV, especialment el HPV16, amb altres tipus de càncer com el d'orofaringe (Herrero et al., 2003). Els gèneres de HPV β, γ, μ, i ν s'associen a patologies de la pell, com ara condilomes, però també s'han detectat de forma habitual en individus sans, suggerint que alguns genotipus de HPV serien part del viroma natural de la pell (Wylie et al., 2014).

Estudis de metagenòmica realitzats en sèrum humà de persones sanes també han descrit la seva presència reforçant la idea que podrien ser considerats virus comensals (Ma et al., 2014; Moustafa et al., 2017). Es creu que la infecció epitelial per part dels HPV succeeix a edats primerenques el nombre d'infeccions amb diferents genotips incrementa amb l'edat. Alguns autors han suggerit un potencial rol d'aquests virus en el desenvolupament de determinats càncers de pell. S'ha observat que determinades lesions de la pell, considerades precursores de càncer com ara la keratinosis actínica, es correlacionen amb una alta carrega viral de HPVβ. L'any 2006 va aparèixer la primera vacuna per a HPV incloent els genotips més freqüentment associats a càncer: HPV16 i 18. Aquesta vacuna s'ha anat ampliant fins a incloure 9 genotips diferents del virus i protegint no només contra aquells associats a càncer sinó també contra alguns genotips associats a condilomes. Alguns autors temen l'eradicació dels genotips inclosos en la vacuna pugui alterar el patró de virus circulants per la població i fomentar l'aparició de nous genotips. És per això que la identificació dels genotips circulants és important per al desenvolupament de programes terapèutics i preventius per als HPV. L'estudi dels HPV a nivell ambiental pot donar informació epidemiològica dels virus circulants. Els HPV poden excretar-se a través de la femta (Di Bonito et al., 2015), de l'orina (Santiago-Rodriguez et al., 2015) o bé a partir de la descamació de la pell (Foulongne et al., 2012) i arribar a les plantes de tractament d'aigua. Recentment s'ha descrit la presència de HPV16 i altres HPV a aigua residual urbana (Bibby and Peccia, 2013a; Cantalupo et al., 2011; La Rosa et al.,

2013b; Symonds et al., 2009a). Aquests virus són molt estables a l'ambient i segueixen sent infecciosos en superfícies contaminades (Ryndock and Meyers, 2014). S'ha observat que són altament resistents a la dessecació i als tractaments esterilitzants emprats en l'entorn mèdic (Meyers et al., 2014; Roden et al., 1997). Mentre que la transmissió sexual dels HPV està ben documentada, la transmissió per via fecal-oral a través del consum d'aliments/aigua contaminats podria ser important. Alguns autors han senyalat el potencial paper dels HPV en el desenvolupament de càncers colorectals (CCR) (Baandrup et al., 2014; Damin et al., 2013; Moreas et al., 2014; Picanço-Junior et al., 2014) però aquest possible paper segueix sent polèmic (Chen et al., 2015). La implicació dels HPV en altres tipus de càncers no genitals ni epitelials també ha estat senyalada (Fratini et al., 2014; Ryndock and Meyers, 2014).

1.6.9. Parvoviridae

La família *Parvoviridae* està constituïda per virus d'unes 5 kbs amb genoma ADN de cadena senzilla. La família *Parvoviridae* es subdivideix en dues subfamílies, conegudes com *Parvovirinae* (infectant principalment vertebrats) i la *Densovirinae* (que infecta artròpodes). Dintre la subfamília *Parvovirinae* existeixen 8 gèneres diferents: *Amdoparvovirus, Aveparvovirus, Bocaparvovirus, Copiparvovirus, Dependoparvovirus, Erythroparvovirus, Protoparvovirus* and *Tetraparvovirus* (Cotmore et al., 2014). Fins fa relativament poc només el Parvovirus B19 era l'únic membre de la família considerat patògen per als éssers humans causant eritema infecciós en persones sanes i immunocompetents o produint, en persones immunocompromeses, anèmia crònica o aplàsia pura de glòbuls vermells (Heegaard and Brown, 2002). Altres parvovirus com ara els Bocavirus Humans (HBoV1-4) o els Bufavirus humans (HBuV) s'han associat a patologies respiratòries o gastrointestinals (Albuquerque et al., 2007; Väisänen et al., 2014a; Wang et al., 2011). Molts autors han descrit la presència de parvovirus en femta on s'han detectat altres virus patògens suggerint que serien virus comensals o agents co-infectants de gastroenteritis agudes (Allander et al.,

2005; Arthur et al., 2009; Han et al., 2009; Jones et al., 2005; Amit Kapoor et al., 2009; Phan et al., 2012). Malgrat això la detecció de parvovirus també s'ha fet en individus sense cap tipus de manifestació clínica (Flotte and Berns, 2005; Moustafa et al., 2017) o s'ha associat a problemes de contaminació en reactius (Naccache et al., 2013). Amb l'aplicació de tècniques de seqüenciació massiva s'han detectat noves espècies tant en éssers humans com en animals (Figura 13) demostrant la gran diversitat d'aquesta família de virus (Bodewes et al., 2014a; Handley et al., 2012; Phan et al., 2016; Shan et al., 2011).

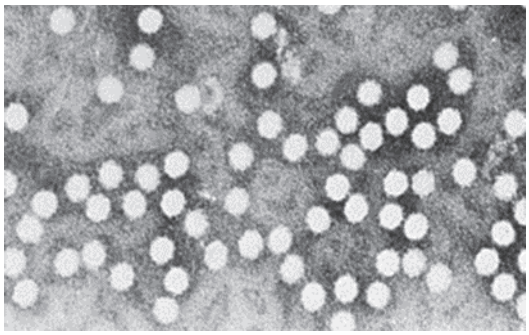


Figura 13. Microscopia electrònica del parvovirus H-1 de rata. Font: CDC

Els parvovirus humans s'han detectat en mostres clíniques de diferents àrees geogràfiques, principalment a femta o mostres de sèrum (Norja et al., 2007; Väisänen et al., 2014). Una de les possibles vies de transmissió per als parvovirus és la fecal-oral. En aigua residual s'ha detectat una alta prevalença de parvovirus superior al 80% (Blinkova et al., 2009). La seva presència també ha estat estudiada en mostres d'aigua de riu sent un 40% positives per mètodes moleculars (Hamza et al., 2009). La seva resistència a baixos pH i altes temperatures (Baylis et al., 2013; Nath Srivastava and Lund, 1980) juntament amb la seva alta prevalença ambiental semblen reforçar una possible via de transmissió fecal-oral, especialment per als Bocavirus Humans 2-4 que es detecten principalment en femta (Kapoor et al., 2010). Finalment, la seva detecció en carn de pollastre, porc i vedella al punt de venta (Peretti et al., 2015; Zhang et al., 2014) representa un interessant descobriment que obre una nova

via d'investigació sobre el possible paper zoonòtic d'aquests virus a través del consum de carn infectada.

1.6.10. Circoviridae

La família *Circoviridae* infecta a mamífers i espècies d'aus. Els circovirus tenen un genoma circular de cadena senzilla de ADN d'unes 1,8 a 3,8 kbs, encapsidats en una càpsida icosaèdrica sense embolcall d'uns 15 a 20 nm de diàmetre. Actualment la família *Circoviridae* conté dos gèneres reconeguts: *Circovirus i Gyrovirus* (Figura 14).

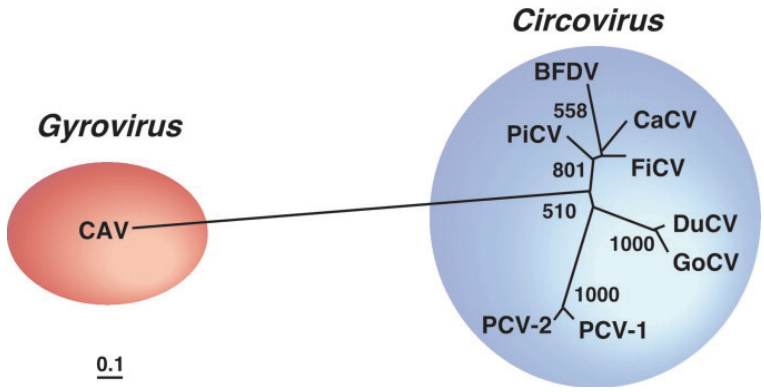


Figure 14. Arbre filogenètic de la familia *Circoviridae* incloent els dos gèneres *Gyrovirus* I *Circovirus* ("Family – Circoviridae," 2012)

Els membres del gènere *Circovirus* tenen el genoma més petit que infecta a mamífers amb una mida de 1,7 a 2 kbs. El seu genoma té dues pautes de lectura oberta que codifiquen per una proteïna de la càpsida (cap) i una proteïna encarregada de replicar el genoma víric (rep). Existeix una regió intergènica entre els gens *rep* i *cap* que conté l'origen de replicació del virus (*ori*). El gènere *Gyrovirus* té un genoma més gran d'unes 2,3 kb. A més el seu genoma està organitzat en tres ORFs parcialment solapats. L'ORF1 codifica per la proteïna de la càpsida mentre que els ORF2 i ORF3 codifiquen per les proteïnes no estructurals VP2 i VP3 del virus (Freed and Martin, 2013).

Els membres de la família *Circoviridae* són coneguts per causar patologia severa en aus i porcs on representen un problema econòmic important. El Circovirus Porci 2 (PCV-2) s'ha detectat en les femtes d'un 5% dels adults americans, presumiblement pel consum de carn de porc infectada (Li et al., 2010).

Existeix un tercer gènere proposat dins la família *Circoviridae* recentment descrit en femtes humanes i de ximpanzés anomenat *Cyclovirus*. Aquests virus han demostrat tenir una alta prevalença en el teixit muscular d'animals de granja, com ara pollastres, vaques i ovelles. La seva detecció també s'ha produït en femta humana on s'han vist divergències de seqüencia importants suggerint que les especies detectades són humanes i no d'origen dietari (Li et al., 2010). El paper d'aquests circovirus detectats en femta humana amb infeccions entèriques encara resulta una incògnita. Recentment s'han detectat seqüències relacionades amb el gènere *Cyclovirus* en femtes de pacients amb afectacions agudes del sistema nerviós central suggerint una possible transmissió per via fecal-oral (Tan et al., 2013).

1.6.11. Picobirnaviridae

Els picobirnavirus (PBV) són virus esfèrics, sense embolcall d'aproximadament uns 33-41 nm de diàmetre i càpsida icosaèdrica amb un genoma de cadena doble d'ARN bisegmentat (Freed and Martin, 2013). Taxonòmicament, la família *Picobirnaviridae* conté un únic gènere *Picobirnavirus*. El seu descobriment data de 1988, quan van ser aïllats a partir de femta humana durant un brot de gastroenteritis aguda al Brasil. Es van descobrir de forma totalment casual mentre les mostres eren testades per rotavirus mitjançant electroforesi en gel de poliacrilamida (Pereira et al., 1988). L'any 2005 es va seqüenciar el primer genoma complet d'un picobirnavirus humà a partir de les femtes d'un nen de 6 anys amb gastroenteritis (Wakuda et al., 2005). El segment 1 del genoma conté dos ORFs. L'ORF1 codifica per una proteïna hidrofílica de funció desconeguda mentre que l'ORF2 codifica per la proteïna de la càpsida. El segon segment del

genoma conté un únic ORF que codifica per la ARN Polimerasa ARN depenent del virus (**Figura 15**).

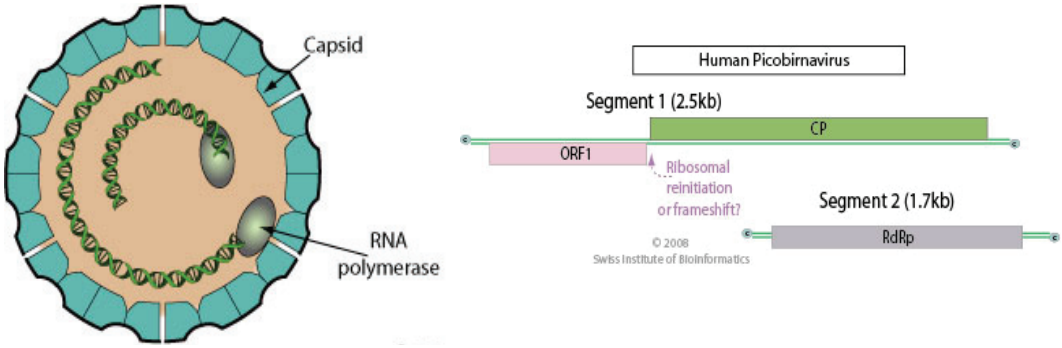


Figura 15. Representació esquemàtica d'un virió de picobirnavirus humà i el seu característic genoma bipartit. Imatge obtinguda de Viral Zone: http://viralzone.expasy.org

Els PBV infecten un ampli rang d'hostes entre els quals trobem els humans, mamífers i aus (Ganesh et al., 2014). Els picobirnavirus humans i animals s'han associat a gastroenteritis i diarrea aquosa, on s'han detectat com a únic agent etiològic possible (Bhattacharya et al., 2007; Browning et al., 1991; Cascio et al., 1996) o en presencia d'altres patògens (Alfieri et al., 1994; Smits et al., 2014). La seva presència en animals i persones sanes ha posat en dubte la seva patogenicitat (Kapusinszky et al., 2012).

L'excreció del virus es fa a través de la femta (Pereira et al., 1988) però no es descarten altres possibles vies d'excreció, com ara la respiratòria, ja que el virus també s'ha detectat al tracte respiratori (Smits et al., 2012). Els Picobirnavirus humans tenen una distribució mundial i s'han detectat amb elevada freqüència a les aigües residuals d'entrada als Estats Units (Symonds et al., 2009b) on totes les mostres d'aigua residual eren positives (12/12). Estudis a Alemanya mostren uns valors més modestos amb una menor prevalença en aigua de riu (23/108) i aigua residual (3/12) presentant unes concentracions de $4{,}0x10^{2}$-$2{,}9x10^{4}$ GC/L i $1{,}0x10^{3}$- $1{,}0x10^{6}$ CG/L respectivament (Hamza et al., 2011). Factors com la manca

de models animals i línies cel·lulars on replicar el virus, juntament al baix nombre de seqüències disponible a les bases de dades, dificulten l'estudi d'aquesta família de virus.

2. Mètodes d'estudi de virus en l'ambient: concentració, extracció i detecció

L'ús de virus com a indicadors es veu dificultat per l'efecte dilució que experimenten en alliberar-se a l'ambient, fent que la seva concentració sigui baixa. És per aquest motiu que es fa necessari un pas de concentració a partir de volums grans de mostra. Aquest volum variarà oscil·lant entre els mil·lilitres a centenars de litres en funció de la contaminació fecal esperada. Existeix un ampli ventall de mètodes de concentració de virus en mostres ambientals, tot i que cap d'ells és considerat el mètode de referència. Segons Block and Schwartzbrod (1989) les característiques desitjables per a qualsevol mètode de concentració es presenten a la següent llista:

- Tècnicament fàcil
- Ràpid processament
- Bona recuperació
- Volums representatius
- Econòmic
- Reproduïble
- Aplicable a diferents virus
- Processar diverses mostres
- Útil en diferents matrius d'aigua

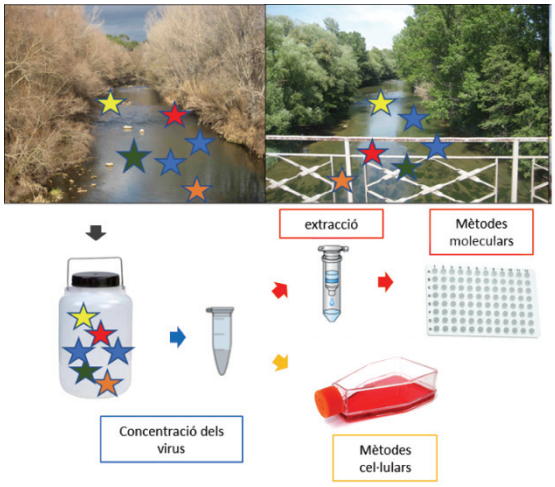


Figura 16. Efecte dilució dels virus en l'ambient. Tècniques de concentració i extracció per al seu estudi.

2.1. Mètodes de concentració de partícules víriques a partir de mostres d'aigua mediambientals

Malgrat l'existència de diferents mètodes de concentració, cap dels mètodes aconsegueix reunir tots aquests requisits. Existeix molta recerca activa per desenvolupar noves metodologies que permetin superar les limitacions dels mètodes actuals i/o desenvolupar-ne de nous. Tradicionalment, els mètodes de concentració de virus en aigua s'han classificat com a primaris o secundaris (o de reconcentració).

Els sistemes de concentració primaris són aquells que permeten passar de grans volums d'aigua fins a un volum final d'aproximadament 1L. Els mètodes de concentració secundari o reconcentració són aquells que permeten recuperar els virus presents en mostres de menys d'1L a volums de pocs mil·lilitres. Malgrat aquesta distinció teòrica, a la pràctica, no sempre resulta fàcil distingir quins mètodes són primaris o secundaris. Històricament, els mètodes de concentració més utilitzats han estat els basats en dos passos d'adsorció- elució (Wallis and Melnick, 1967), també coneguts com a mètodes VIRADEL (*VIRal ADsorption and ELution*). En aquests, s'aconsegueix la concentració de partícules víriques de la mostra per adsorció electrostàtica i posterior elució amb tampons alcalins (generalment utilitzant extracte de carn, solucions salines, glicina, etc.) a partir de filtres, cartutxos o partícules que retenen els virus. Aquests mètodes permeten filtrar grans volums d'aigua i eluir els virus en un volum relativament petit. Utilitzar grans volums d'aigua implica una co-concentració de molècules inhibidores que poden afectar la seva posterior detecció (Gibson et al., 2012; Ijzerman et al., 1997).

Els diferents virus entèrics presenten un ampli ventall de proteïnes de la càpisda diferents en funció de l'espècie; totes elles presentant diversitat de mides i càrregues. El diàmetre dels virus oscil·la entre els 30nm (enterovirus) i els 100nm (adenovirus), amb rangs de punt isoelèctric que oscil·len entre els 2,8 de l'HAV als 8,0 dels RoV (Michen and Graule, 2010). Tot i aquestes diferències de mida i

forma, els virus presenten en el medi aquós una càrrega electronegativa, fent necessari el condicionament de les mostres o bé de la superfície dels filtres per facilitar-ne l'absorció.

Entre els mètodes de concentració més utilitzats basats en adsorció-elució cal destacar:

• Filtres electronegatius (Farrah et al., 1976). Composats per una barreja d'acetat de cel·lulosa, nitrat de cel·lulosa o llana de vidre. Els virus en aigua tenen tendència a unir-se a filtres electronegatius en presència de cations multivalents (Al^{3+}, Mg^{2+}) sota condicions àcides (pH=3,5), ja que tenen una superfície de càrrega electropositiva en aquestes condicions. Els virus retinguts són eluits utilitzant glicina a pH bàsic. El filtrat és posteriorment neutralitzat fent servir H_2SO4 i tampó TE 100X.

• Filtres electropositius (Sobsey and Glass, 1980). Sota condicions de pH pròximes a la neutralitat la majoria dels virus presenten una carrega isoelèctrica neutre o lleugerament negativa. Això permet la seva adhesió damunt de filtres electropositius sense necessitat d'ajustar el pH o afegir cations multivalents (Al^{3+}, Mg^{2+}) a la mostra. El mètode es basa en fer passar les mostres d'aigua per el filtre electropositiu. A continuació s'elueixen els virus adsorbits mitjançant una solució bàsica que modifica el punt isoelèctric del virus. Aquesta solució sol basar-se en una solució amortidora d'extracte de vedella (*beef extract buffer*) amb glicina 0,05M a pH de 9,5. Ràpidament l'eluït es neutralitzat amb una solució àcida de HCl 1N per tal d'evitar l'inactivació de les partícules víriques. L'únic mètode aprovat per la USEPA (*United States Environmental Protection Agency*) per recollir virus a partir de mostres d'aigua de distribució es basa en l'ús de filtres electropositius 1MDS (CUNO, Meriden, CT, USA). Recentment han sortit al mercat els filtres NanoCeram® (Argonide, Sanford, FL, USA) i els filtres Virocap (Scientific Methods, Inc., Granger, IN, USA), que milloren algunes de les limitacions dels filtres electropositius existents.

• Llana de vidre (Vilaginès et al., 1993). La llana de vidre empaquetada a una determinada densitat dins d'una columna actua com a material absorvent de partícules víriques. La llana de vidre pot utilitzar-se sola o bé unida a un agent d'unió i revestida d'una capa d'oli mineral. Aquest mètode presenta a la seva superfície llocs d'unió hidrofòbics (capa d'oli) i punts d'unió electropositius. Quan una suspensió que conté virus passa pel porus generat amb el material empaquetat, la superfície de fibra és capaç d'atrapar i retenir les partícules víriques carregades negativament a un pH prop de la neutralitat (pH pròxim a 7). S'ha emprat bàsicament per concentrar virus provinents de mostres d' origen ambiental de gran volum.

• *Monolithic Affinity Filtration* (MAF) (Kunze et al., 2015; Pei et al., 2012). Les columnes MAF permeten la concentració de partícules víriques mitjançant l'adsorció-elució dels virus a un disc compost per un polímer que té radicals epoxy. Per generar les columnes MAF es fa servir un compost, el poliglicerol-3-glicidil, que té capacitat d'autopolimeritzar, en presència de toluè i terc-butil metil èter que actuen com a agents generadors de porus. Les columnes es netegen amb metanol i la superfície monolítica ha d'activar-se amb un tractament d'àcid sulfúric (0,5M) a 60ºC durant 1 hora. Aquestes columnes actives es dipositen dintre d'un sistema columnar pel que va entrant la mostra d'aigua pre-condicionada a pH 3,0. A pH àcid els virus tenen càrrega isoelèctrica positiva de manera que queden adsorbits a la superfície de les columnes monolítiques per interaccions iòniques i hidrofòbiques. Els virus s'elueixen de la columna fent servir 20 mL de solució amortidora amb altes concentracions de sal (1,5 M NaCl, 0,05 M HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) a pH 7).

També hi ha un mètode basat en floculació orgànica en un únic pas:

• Floculació directa amb llet descremada (Calgua et al., 2013b, 2013c, 2008). Es basa en la unió directa dels virus a la llet descremada prefloculada. Primerament

cal elaborar una solució al 1% de llet descremada en aigua de mar artificial i ajustar el pH fins a 3,5 afegint HCl 1N. La conductivitat i pH de la mostra han de ser condicionades de la mateixa manera. Les partícules víriques s'adheriran als flòculs i precipitaran. El mètode no requereix l'utilització de filtres ni bombes de pressió, sovint utilitzades en els mètodes de concentració de virus, reduint així l'equipament necessari per a dur a terme l'anàlisi. És un mètode en un sòl pas que permet obtenir un concentrat víric sense necessitat d'aplicar un mètode de concentració secundari i aplicable a diferents matrius (aigua de mar, riu, aigua residual o de pou).

A més, existeixen també mètodes basats en tècniques d'ultrafiltració. Aquests mètodes es basen en la retenció, en funció de la mida, de les partícules víriques i altres microorganismes i no pas per la seva càrrega isoelèctrica. La mostra d'aigua es fa passar a través de les membranes o de cartutxos que retenen aquelles partícules amb pesos moleculars entre 30-100kDa. Degut al tamany del porus, l'aigua i altres molècules més petites, passen a través mentre que els virus queden retinguts.

• Ultrafiltració per membranes (Belfort et al., 1982; Divizia et al., 1989). Es basa en la utilització de membranes d' ultrafiltració per concentrar i recuperar microorganismes presents en mostres d'aigua. La mostra d'aigua passa una única vegada a través del filtre. Un cop finalitza el procés de filtració el retingut és renta i resuspèn amb una solució amortidora per obtenir les partícules víriques. Pot aplicar-se com a mètode de concentració primàri utilitzant cartutxos d'ultrafiltració o com a mètode de concentració secundari utilitzant filtres tipus Amicon.

• Filtració i ultrafiltració tangencial (TFF) (Wommack et al., 2010). A diferència de la filtració convencional, en la TFF, la mostra circula de forma tangencial respecte a la membrana filtradora, evitant el procés d'obturació. El sistema utilitza una bomba peristàltica d'alta pressió que fa passar la mostra pel filtre de flux

tangencial. La fracció retinguda, la qual és incapaç de passar pel filtre, es farà recircular repetidament afavorint l'obtenció del concentrat en un menor volum.

• Vortex flow filtration (VFF) (Paul et al., 1991). És una tecnologia basada en filtració que utiliza el corrent generat pels vòrtex de Taylor com a força filtradora. Els vòrtex de Taylor es generen en col·locar un cilindre filtrador rotador dins d'un cilindre de diàmetre major. La mostra a filtrar s'injecta a una pressió controlada i es fa circular entre els dos cilindres, forçant l'entrada de les molècules dins el segon cilindre, des d'on es guardaran o descartaran en funció de la mida desitjada. El retingut que conté els virus retorna al sistema d'injecció i es recircula diverses vegades per a concentrar la mostra en un menor volum. El corrent generat pels vòrtex impedeix la colmatació del sistema de filtració.

Entre els principals mètodes de concentració secundaris trobem:

• Polyethilene glycol precipitation (PEG) (Lewis and Metcalf, 1988). En aquesta tècnica els virus en suspensió són precipitats amb PEG. El PEG és un agent higroscòpic que remou l'aigua produint la insolubilització i subseqüent precipitació dels virus. La precipitació i agregació dels virus permet aplicar centrifugacions lleugeres que d'altra manera no serien eficients per a concentrar virus.

• Floculació orgànica (Katzenelson et al., 1976). És un mètode de concentració principalment secundari. Aquest procés es basa en l'acidificació de la mostra per sota del punt isoelèctric de les proteïnes (pH 3,5), provocant l'adsorció dels virus a la matriu orgànica (extracte de carn, llet descremada...). Els flòculs formats són centrifugats i les partícules víriques resuspeses en un buffer a pH neutre. Existeix una variant que utilitza clorur fèrric al 25 mM per com a coadjuvant del procés de floculació (Payment et al., 1984).

• Concentració per adsorció-precipitació amb hidròxid d'alumini (Farrah and Preston, 1985). Es creu que és deguda a interaccions electrostàtiques entre la superfície del virus, que té carrega negativa, i la superfície de càrrega positiva de

l'hidròxid d'alumini. Els virus queden adsorbits en un precipitat d'hidròxid d'alumini que és afegit a la mostra o que es forma en ella gràcies a l'addició de carbonat de sodi o hidròxid de sodi i la presència d'una sal soluble d'alumini. Un cop s'ha permès l'adsorció dels virus al precipitat aquest es recull per filtració o centrifugació. Pot utilitzar-se directament el precipitat per fer anàlisi d'infectivitat al laboratori, o bé, eluir del precipitat els virus fent servir un tampó alcalí o una solució proteica.

• Ultracentrifugació (Pina et al., 1998a). Generalment es tracta d'un mètode de concentració secundari que pot fer-se servir com a mètode de concentració primari en determinades matrius molt contaminades com ara l'aigua residual. La solució que conté els virus eluïts procedents del procés de concentració primari o la mostra a estudiar és sotmesa a un procés d'ultracentrifugació d'unes 220.000xg 1 h a 4°C. El sediment és resuspès amb 3,5mL de tampó glicina (pH 9,5) durant 30 min en gel. Seguidament s'afegeix PBS i es separen els sòlids resuspesos amb centrifugació (12.000 xg durant 15 minuts). Retenint el sobrenedant, la mostra es sotmesa a un nou procés d'ultracentrifugació a 220.000xg 1 h a 4°C que ens permetrà concentrar les partícules víriques. El *pellet* víric serà resuspès en 0,1 mL de PBS. Al no dependre ni de la càrrega ni de la mida presenta percentatges de recuperació molt elevats.

A la Taula 3 es fa una valoració qualitativa del grau d'adequació dels mètodes de concentració primaris de virus explicats anteriorment segons diferents variables o el tipus de matriu a estudiar.

Taula 3. Valoració qualitativa del grau d'adequació dels principals mètodes de concentració primaris de virus segons les variables o matrius estudiades a la literatura. Llegenda: verd (apte), groc (no concloent), vermell (no apte), blanc (insuficient informació), (Mat.: materials).

Mètode / Variables o matrius	Pre-condicionament	Bona recuperació	Volum grans (≥100L)	Problemes mat. suspensió	Facilitat d'ús	Simultaneïtat mostres	Ús en terreny	Rapidesa	Disponibilitat	Preu	Viabilitat	Concentració varis patògens	Aigua continental	Aigua de mar	Aigua residual
Filtres electronegatius															
Filtres electropositius															
Filtres Nanoceram®															
Filtres Virocap															
Llana de vidre															
MAF															
Ultrafiltració															
Ultrafiltració tangencial															
Vortex Flow Filtration															
Ultracentrifugació															
Floculació llet descremada															

2.2. Tècniques tradicionals per a l'estudi de virus al ambient

Un cop s'ha produït la concentració de les partícules víriques és necessari un procés de detecció, quantificació i identificació dels virus presents. Tradicionalment s'han utilitzat tècniques basades en el cultiu cel·lular o mètodes moleculars com la PCR o PCR quantitativa (qPCR). Recentment, el desenvolupament de tècniques de seqüenciació massiva ha permès l'estudi de les poblacions víriques des d'una nova perspectiva: la metagenòmica.

2.2.1. Mètodes basats en PCR per a l'estudi de virus

Els mètodes més utilitzats per a l'estudi de virus a l'ambient són els mètodes moleculars. Això es degut a que són mètodes molt sensibles, específics, ràpids i econòmics en comparació amb els mètodes basats en cultiu. La PCR permet detectar la presència de genomes vírics mitjançant l'ús d'encebadors específics que amplifiquen regions conservades del virus. Sovint la concentració d'aquests virus és molt baixa a l'ambient requerint l'aplicació d'una segona PCR amb encebadors més interns anomenada PCR niuada o *nested*-PCR (dos encebadors interns) o semi-niuada o semi-*nested* (un encebador intern). Aquesta segona PCR permet augmentar la sensibilitat del mètode en mostres ambientals. Les PCR niuades de regions variables del virus s'analitzen per a fer estudis de tipificació i filogènia de virus. La tècnica de la PCR quantitativa o a temps real (qPCR o q(RT)-PCR) permet, a més de detectar, quantificar la presència de genomes vírics de forma molt fiable (Girones et al., 2010). Malgrat això, el fet de quantificar directament genomes no permet tenir dades sobre la infectivitat dels virus detectats.

Dels múltiples mètodes existents per extraure àcids nucleics, els basats en kits d'extracció comercials són molt utilitzats. Aquesta extracció es basa en quatre passos: trencament de la càpsida proteica del virus, adsorció dels àcids nucleics a una matriu, rentats per purificar i netejar els àcids nucleics d'altres mol·lecules com proteïnes i restes cel·lulars, i finalment elucció dels àcids nucleics. Els kits comercials més utilitzats es basen en l'ús de columnes de sílice que retenen els àcids nucleics o bé en l'utilització de partícules magnètiques que s'uneixen al ADN.

Malgrat les limitacions dels mètodes moleculars, es pot determinar de forma aproximada si el ADN/ARN detectat correspon a virus infecciosos. Una opció és l'ús d'anticossos marcats magnèticament per aïllar partícules víriques a partir de mostres d'aigua i posteriorment utilitzar tècniques moleculars per a la seva detecció (Abd El Galil et al., 2004; Yang et al., 2011). L'ús de mol·lècules com

l'EMA (*ethidium monoazide*) o el PMA (*propidium monoazide*), capaços d'atravessar membranes lipídiques o nucleocàpsides danyades, s'uneixen covalentment al ADN o ARN en ser exposades a llum ultraviolada. Aquesta unió irreversible inhibeix posteriorment la reacció de PCR, facilitant l'amplificació d'aquells virus que tenen les seves estructures intactes, i per tant, possiblement infectius (Kim et al., 2011; Parshionikar et al., 2010; Quijada et al., 2016). L'utilització de nucleases que eliminin l'ADN/ARN lliure, prèviament a l'extracció dels àcids nucleics, també s'ha utilitzat com a aproximació per estudiar virus viables (Pecson et al., 2009).

2.2.2. Cultiu cel·lular per l'estudi de virus infecciosos

Els mètodes basats en cultiu cel·lular ofereixen informació sobre la presència de virus infecciosos a la mostra. Aquesta informació és especialment útil i important si les dades derivades de l'anàlisi s'han d'utilitzar per a fer estudis de risc. Des d'un punt de vista pràctic es considera que un virus és infecciós quan és capaç d'entrar dins la cèl·lula i replicar-se. L'efecte que genera la infecció viral en la cèl·lula hoste es coneix com a efecte citopàtic (CPE). Aquest pot manifestar-se de diverses formes en funció de la línia cel·lular i el virus que la infecta: generació de cossos d'inclusió, arrodoniment cel·lular, pèrdua d'adhesió al flascó, lísi cel·lular... Entre els mètodes comunament emprats per a estudiar la infectivitat trobem el càlcul de la dosi necessària per infectar un cultiu al 50%, conegut com a TCID50 (*Tissue Culture Infectious Dose* 50%), el recompte de calbes de lisi (PFU) o l'observació de proteïnes virals utilitzant anticossos marcats amb immunofluorescència dins les cèl·lules hoste (FFU). Malgrat els avantatges que ofereixen els mètodes basats en cultiu, existeixen diverses limitacions que afecten al seu ús com a mètode rutinari. En primer lloc, no tots els virus produeixen CPE evident o el temps necessari per a produir-lo és molt llarg. Les mostres ambientals són mostres complexes que poden contenir un ampli ventall de virus i altres molècules que poden afectar la viabilitat de les línies cel·lulars, resultant tòxiques. A més, no tots els virus transmesos per aigua i/o aliments

contaminats disposen d'una línia cel·lular on cultivar-se. L'exemple més clar el trobem amb els norovirus humans, principal agent víric transmès per aigua contaminada, el cultiu del qual és molt laboriós. No ha estat fins l'any 2016 que s'ha aconseguit cultivar-los, requerint de mètodes de cultiu molt costosos i difícils d'aplicar per a l'estudi de mostres ambientals (Ettayebi et al., 2016). Finalment, la necessitat de mà d'obra molt especialitzada i l'elevat preu derivat del manteniment i propagació de les línies cel·lulars també són inconvenients importants.

Per tal de solventar algunes de les limitacions dels mètodes de cultiu, com són l'absència o dificultat d'interpretació d'alguns CPE o els llargs temps d'incubació per a observar-los, s'han desenvolupat tècniques com la ICC-PCR (Integrated Cell Culture-PCR). Aquesta tècnica combina la rapidesa i sensibilitat de les tècniques moleculars amb informació sobre la viabilitat dels virus presents a la mostra (Greening et al., 2002; Jiang et al., 2004; Reynolds et al., 2001; Shieh et al., 2008). La ICC-q(RT)PCR és una variant de la ICC-PCR que incorpora real-time PCR. De forma simple, es basa en quantificar la mostra abans i després del seu pas pel cultiu cel·lular per a veure un increment en el nombre de genomes vírics (Balkin and Margolin, 2010; Ogorzaly et al., 2013).

2.3. Noves metodologies per a l'estudi de virus

Des del desenvolupament del "mètode dideoxi" creat per Sanger i col·laboradors (Sanger et al., 1977) les tècniques de seqüenciació han canviat de forma considerable. El conjunt d'innovacions en el camp de la seqüènciació han evolucionat de forma molt ràpida des de l'obtenció del primer genoma bacterià a mitjans dels anys 90 (Fleischmann et al., 1995; Fraser et al., 1995), especialment a partir del que s'ha anomenat l'era post-genòmica i la publicació dels genomes d'organismes complexos com ara el dels humans i altres eucariotes (**Figura 17**).

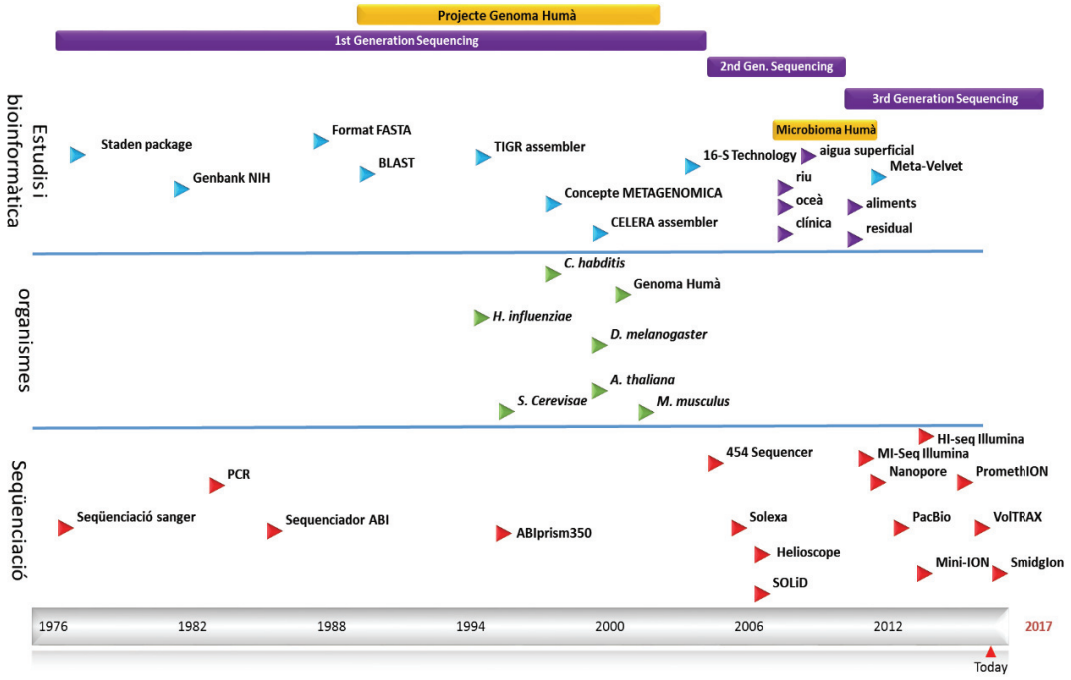


Figura 17. Cronograma representatiu dels principals avenços tecnològics en el camp de la seqüenciació i la bioinformàtica

Les millores tecnològiques han permès reduir de forma significativa els costos de seqüenciació per base, permetent la democratització de les noves tècniques de seqüenciació a la majoria de laboratoris (Shendure and Ji, 2008; Mardis, 2008; Schuster, 2008), malgrat que el cost s'ha transferit a l'emmagatzematge i gestió de la ingent quantitat de dades que es generen i a la posterior i més laboriosa anàlisi d'aquestes dades (Sboner et al., 2011) ; això inclouria la neteja i

l'ensamblat de les traces de seqüenciació, però també la interpretació dels resultats sobre mostres de matrius i dissenys experimentals més complexos. En el nostre camp el *boom* en la producció de dades de seqüència s'exemplifica que en el moment de redactar aquesta tesi hi hagi més de 71.628 seqüències virals validades a la base de dades NCBI RefSeq, divisió viral (https://www.ncbi.nlm.nih.gov/genome/viruses/).

2.3.1. Del genoma al metagenoma

La metagenòmica representa l'estudi de tot el material genètic disponible en una mostra, sovint de tipus ambiental, en la que pot aparèixer una barreja d'espècies diverses. El terme va ser introduït per primera vegada per Handelsman i col·laboradors (1998), en un estudi en que es descrivia la composició microbiana del sòl. El seu treball va demostrar que la diversitat microbiana era molt més gran que l'observada a partir de mètodes basats en el cultiu. Més endavant es va utilitzar la seqüència de l'ARN ribosomal 16S com a marcador i eina filogenètica en estudis de metagenòmica enfocats al món dels procariotes (Handelsman, 2004). Aquest marcador s'ha aplicat de forma universal per l'anàlisi de metagenomes bacterians en una gran diversitat d'ambients (Jovel et al., 2016; Sinclair et al., 2015), tot i les seves limitacions (Poretsky et al., 2014). En el cas dels virus, no existeix un marcador universal tipus ARNr-16S bacterià o ARNr-18S eucariòtic, de manera que l'estudi dels metagenomes vírics, també anomenats viromes, es troba limitat per la seqüenciació de tot l'ADN/ARN genòmic present en les mostres. Això requereix aplicar tractaments complexos per tal de reduir la presència d'ADN/ARN d'origen no víric (**Figura 18**). Algunes opcions per aconseguir-ho inclouen: l'ús de mètodes de concentració/extracció que afavoreixin l'enriquiment de partícules virals en la mostra; afegir nucleases per eliminar ADN/ARN lliure, probablement d'origen bacterià i/o eucariòtic; el filtrat de la mostra per eliminar bacteris; el tractament dels concentrats amb cloroform/butanol per eliminar bacteris, però també inhibidors dels enzims emprats en els protocols d'amplificació i seqüenciació; utilitzar tècniques

immunològiques de precipitació per anticossos per capturar les partícules víriques; o l'ús de columnes de poliA per capturar determinats grups de virus, entre d'altres (Hall et al., 2014; Hjelmsø et al., 2017; Oude Munnink et al., 2013).

La metagenòmica aplicada a virus es troba limitada per les baixes quantitats d'ADN/ARN presents a les mostres, en especial aquelles que són d'origen ambiental com les provinents d'aigua i/o aliments. Per obtenir prou ADN per assolir els mínims necessaris per moltes de les tècniques de seqüenciació modernes, sovint cal efectuar una pre-amplificació dels àcids nucleics presents a la mostra mitjançant una PCR amb encebadors aleatoris o "*random primers*" (Reyes and Kim, 1991; Wang et al., 2003, 2002). Els mètodes més utilitzats es basen en l'amplificació de desplaçament múltiple, de l'anglès *Multiple Displacement Amplification* o MDA, o bé en l'amplificació independent de seqüència basada en un únic encebador, de l'anglès *sequence-independent single-primer amplification* o SISPA (Edwards and Rohwer, 2005; Hutchison et al., 2005; Spits et al., 2006).

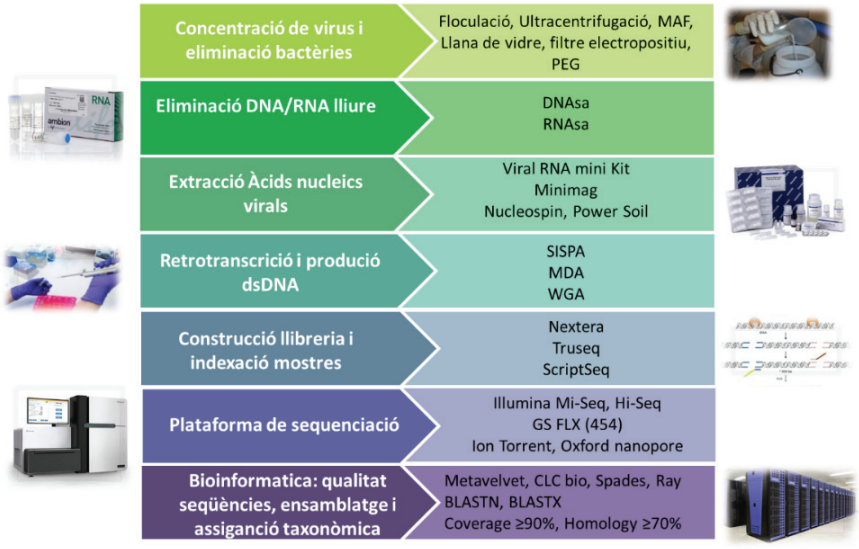


Figura 18. Diagrama de pasos necessaris per estudiar el viroma d'una mostra mitjançant NGS

L'utilització de mètodes d'amplificació independents de seqüència permet detectar, en un únic assaig, virus nous i variants altament divergents dels ja descrits (Bodewes et al., 2014b).

L'aplicació de NGS al món de la virologia ambiental ha fet possible l'estudi dels viromes d'oceans (Hurwitz et al., 2013), llacs (Djikeng et al., 2009), aigua residual (Cantalupo et al., 2011), aigua regenerada (Rosario et al., 2009b) o aliments (Aw et al., 2016). Aquestes tècniques també s'han aplicat a mostres clíniques per a les que es desconeixia l'etiologia de l'agent causal (Greninger et al., 2015; Moore et al., 2015; Sauvage et al., 2016; Tan et al., 2013). Aquestes tècniques han popularitzat i afavorit una explosió dels estudis basats en seqüenciació, produint un increment important del nombre de seqüències dipositades a les bases de dades, així com el desenvolupament de bases de dades específiques, com per exemple MetaVir (http://metavir-meb.univ-bpclermont.fr/) o altres recursos metagenòmics (https://www.ebi.ac.uk/metagenomics/). També s'han desenvolupat diverses plataformes que permeten comparar metagenomes de diferents indrets, per exemple MG-Rast (http://metagenomics.anl.gov/)(Meyer et al., 2008).

Prèviament a l'inici de la present tesi doctoral, el nombre de publicacions centrades en l'estudi del viromes en matrius ambientals era reduït (**Taula 4**).

Taula 4. Estudis que apliquen tècniques de seqüenciació massiva en mostres ambientals fins l'any 2013.

Títol	Autor i Any	Matriu	Virus detectats
A bacterial metapopulation adapts locally to phage predation despite global dispersal	Kunin et al., 2008	Sediments d'estuari	Bacteriòfags
Metagenomic analysis of viruses in reclaimed water	Rosario et al., 2009	Aigua regenerada	Animals: *Circoviridae, Nodaviridae, Picronaviridae, Dicistroviridae*
Census of the Viral Metagenome within an Activated sludge Microbial Assemblage	Parsley et al., 2010	Fangs actius	*Myoviridae*, *Siphoviridae*, *Podoviridae*, Bacteriòfags
Raw sewage harbors diverse viral populations	Cantalupo et al., 2011	Aigua residual	Animal/humà: *Adenoviridae, Astroviridae, Caliciviridae, Papillomaviridae, Parvoviridae, Picobirnaviridae, Picornaviridae, Polyomaviridae* Bacteriòfags
Viral Metagenome Analysis to Guide Human Pathogen Monitoring in Environmental Samples	Bibby et al., 2011	Biosòlids	Bacteriofags, Animal/humà*: Herpesviridae, Coronaviridae, Picornaviridae, Adenoviridae, Flaviviridae, Circoviridae*
High Variety of Known and New RNA and DNA Viruses of Diverse Origins in Untreated Sewage	Ng et al., 2012	Aigua residual	Animal/humà*: Adenoviridae, Astroviridae, Caliciviridae, Hepeviridae, Parvoviridae, Picornaviridae, Picobirnaviridae, Reoviridae*
Identification of Viral Pathogen Diversity in Sewage Sludge by Metagenome Analysis	Bibby and Peccia, 2013	Biosòlids	Animal/humà: *Papillomaviridae, Adenoviridae, Parvoviridae, Circoviridae, Coronaviridae, Parvoviridae, Reoviridae,Caliciviridae, Flaviviridae, Astroviridae, Herpesviridae*

En determinades ocasions pot ser més convenient dirigir la recerca per estudiar amb major profunditat determinades famílies de virus presents en una mostra. L'ús de mètodes d'amplificació dirigits en combinació amb tècniques de seqüenciació masiva es coneix amb el nom de target enrichment o NGS Amplicon Sequencing (Ogorzaly et al., 2015; Wylie et al., 2015).

Els protocols de seqüenciació massiva es basen en tres processos crítics: la preparació de la mostra d'ADN, la inmobilització i la seqüenciació. Generalment el pas de preparació de l'ADN inclou un pas on seqüències conegudes, anomenades adaptadors, s'uneixen als extrems de l'ADN fragmentat. L'ADN de la mostra al que s'han afegit els adaptadors als extrems es coneix com a llibreria de seqüenciació. El marcatge amb adaptadors és necessari per la fase d'immobilització de la llibreria sobre una fase sòlida on es generen rèpliques en paral·lel de cada fragment, conegudes com a clústers o colònies de la polimerasa (anomenades sovint amb el terme "polònies"), sobre les que es duu a terme el procés de seqüenciació.

Les diferents plataformes disposen de diferents sistemes òptics o electrònics que monitoritzen les reaccions moleculars de síntesi de cadena complementària a cadascun dels clústers de la polimerasa. Un cop processades les imatges o els impulsos els aparells permeten obtenir la composició nucleotídica simultàniament per totes les seqüències de l'experiment de seqüenciació (el que es defineix en anglés com a "sequencing run"). Algunes de les plataformes detecten la incorporació de nucleòtids en cada cicle d'amplificació utilitzant nucleòtids marcats amb fluoròfors, altres detecten canvis de pH digitalment (IonTorrent) o bé químicament (piroseqüenciació 454). Per tancar cada cicle, després de cada incorporació s'apliquen rentats i s'enregistra el senyal molecular corresponent al nucleòtid marcat mitjançant el sistema de detecció específic de cada tècnica.

2.3.2 Tecnologies de seqüenciació

En el moment d'escriure aquesta tesi, Illumina és la principal plataforma de seqüenciació massiva, ocupant aproximadament el 80-85% del mercat. Altres tecnologies estan en desenvolupament, com ara el mètode PacBio o l'Oxford nanopore, ambdues metodologies de seqüenciació sense amplificació (el que hom anomena "*single-molecule sequencing*"), o bé són minoritàries o molt específiques com Ion Torrent, o bé estan en declivi com el 454 de Roche i la tècnica de SOLID.

2.3.2.1. Genome sequencer GS FLX (454)

La preparació de la mostra inclou la fragmentació a l'atzar del ADN de doble cadena en fragments de 300 a 800 parells de bases (Margulies et al., 2005). Els extrems de ADN són reparats per enzims que afegeixen una cua de poliA als extrems 3' als que s'afegiran els adaptadors A i B. L'adaptador B està marcat amb biotina, que en ser detectat per estreptavidina permetrà la purificació dels fragments que contenen els dos adaptadors i els fragments amb quimeres d'homopolímers A o B seran descartats. Els fragments seran capturats per un únic "*bead*" o boleta i amplificats mitjançant el sistema de PCR en emulsió. Cadascuna de les boletes immobilitzarà milers d'amplificacions clonals del fragment inicial. Aquestes boletes seran posteriorment immobilitzades en un suport sòlid que conté milions de nano-pous en els que tan sols hi pot entrar una única boleta. En aquests pouets tindran lloc les reaccions de seqüenciació, basades en l'ús de nucleòtids marcats luminescentment que són detectats de forma progressiva en cada cicle de seqüenciació (**Figura 19**). La limitació més important d'aquesta tècnica és el baix nombre de *reads*, el seu elevat preu, així com biaixos en la seqüenciació de regions amb repeticions de més de tres nucleòtids-per exemple homopolímers CCCC-que afecten a la taxa d'insercions-delecions.

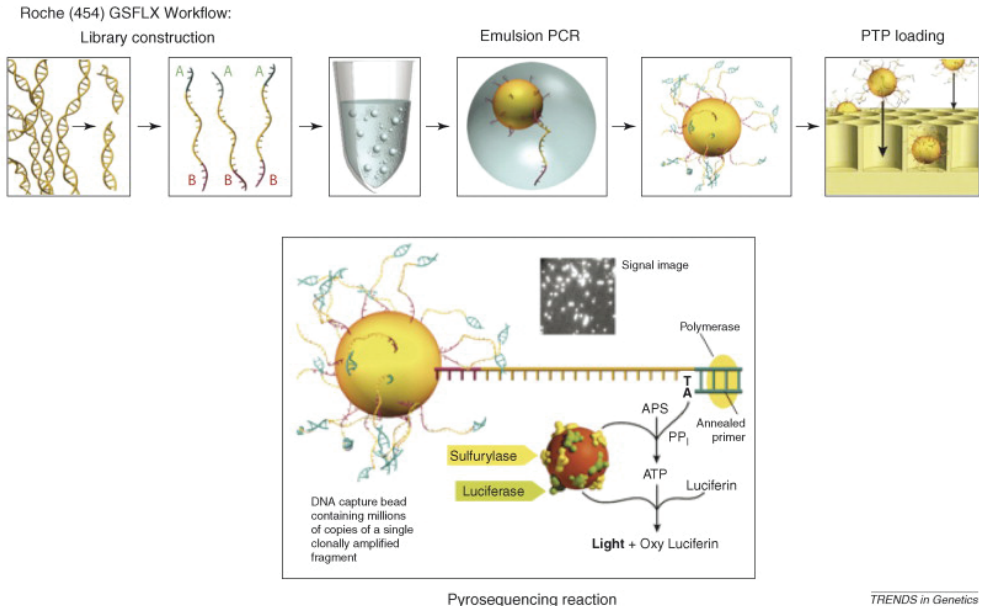


Figura 19. Diagrama de funcionament de la plataforma 454 de Roche (Mardis, 2008)

2.3.2.2. Illumina Sequencer

Per preparar la llibreria es requereix ADN de doble cadena fragmentat, al que s'afegirà en els extrems uns adaptadors i marcadors específics que s'utilitzaran posteriorment per indexar la mostra. Aquest procés de marcatge i indexació es pot fer de maneres diferents en funció de la quantitat de ADN disponible per preparar la llibreria, però també en funció del nombre de mostres que es poden seqüenciar conjuntament (multiplexació). Alguns dels sistemes utilitzats de manera rutinària són els protocols Nextera, Nextera XT i truSeq.

El sistema Nextera es basa en la utilització de transposomes dissenyats per enginyeria genètica que "tagmenta" l'ADN de doble cadena de la mostra. El procés de tagmentació inclou la fragmentació i la col·locació d'adaptadors en un únic pas. Posteriorment aquests adaptadors es faran servir per a fer una reacció de PCR que permetrà indexar les mostres en els dos extrems de l'insert de ADN.

En funció de la quantitat de ADN disponible en la mostra es fa servir el kit de Nextera (50 ng ADN) o el de Nextera XT (1ng).

El sistema truSeq pot utilitzar directament ARN o ADN per a la construcció de la llibreria. En el cas de partir de ARN, aquest es selecciona a partir de les cues de poli-A i es sintetitza una cadena complementària de cDNA fins a obtenir ADN de doble cadena. Es requereix, com a mínim, 1 µg de ADN que es fragmentarà, per sonicació o nebulització, generant extrems roms per l'addició d'exonucleases i enzims de reparació. En els extrems roms s'afegeix una base "A", que servirà per a lligar els adaptadors de seqüenciació. Cadascun dels adaptadors conté una "T" penjant que serà complementaria a la "A", facilitant el lligam de l'adaptador. Un cop lligat en el fragment, l'adaptador conté també la informació referent a l'índex evitant la necessitat d'amplificar per indexar els fragments. Finalment, es duu a terme una reacció de PCR per generar els clústers clonals o "polònies". Existeixen diferents variants dintre del sistema:, truSeq ADN kit (explicat anteriorment, 1 µg de ADN), truSeq ADN nano (requereix 100ng) i truSeq ADN PCR-*free* (evita el pas de PCR final, 1-2 µg de ADN). Els kits truSeq nano i truSeq ADN PCR-*free* semblen oferir millor cobertura (o "*coverage*") per als genomes i produir menor nombre de *gaps*.

La llibreria preparada es carrega i immobilitza a la "*flow cell*" que conté una sèrie de encebadors complementaris als adaptadors. Un cop hibriden la seqüència de la llibreria es fa servir com a motlle per a construir una segona cadena complementària. A continuació es generaran clústers clonals a la superfície de la "*flow cell*" fent servir un procés conegut com a "*bridge* PCR"(Figura 20).

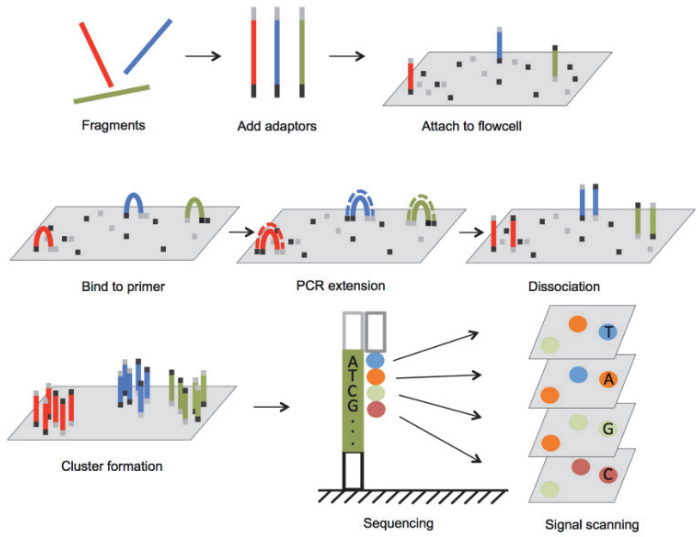


Figura 20. Diagrama de representació del bridge amplification per generar clústers clonals a la plataforma Illumina (Lu et al., 2016).

Aquesta amplificació genera els clústers necessaris per a la correcta lectura de les seqüències. El mètode de seqüenciació per síntesi utilitza 4 dNTPs terminadors diferents marcats amb 4 flouròfors específics per cada nucleòtid. Després de cada reacció d'incorporació, els clústers són monitoritzats en paral·lel obtenint-se la base corresponent a cada posició dins la seqüència. Abans de prosseguir al següent cicle, en aquells fragments que han incorporat un nucleòtid terminador s'elimina, permetent l'addició del següent nucleòtid. La lectura seqüencial d'un sol nucleòtid per cicle assegura una baixa taxa d'error davant homopolímers. Per contra, el fet que existeixi un decalatge en l'adició de nucleòtids dins del clúster al perdre la sincronia les polimerases que sintetitzen la cadena complementària, fa que el senyal vagi disminuint la seva qualitat al llarg dels cicles; això implica que la tecnologia d'Illumina sigui molt bona per a seqüències curtes, sobre uns 300 parells de bases, però no tan apta per a seqüències llargues.

Illumina disposa, pels seus usuaris, d'una gran diversitat d'opcions de seqüenciació que s'ajusten a les necessitats de cadascú en funció de l'aplicació:

Mini-Seq (8 Gbp, 25 $x10^6$ reads, 2x150 paired-ends), Mi-Seq (15Gbp, $25x10^6$ reads, 2x300 paired-ends) o Hi-Seq (1500Gbp, $5x10^9$ reads, 2x150 paired-ends).

2.3.2.3. PacBio

La plataforma Pac Bio RS (Pacific BioSciences) permet seqüenciar una molècula sencera en temps real. Per a preparar les llibreries el ADN es fragmenta, repara i se li afegeix una "A" penjant que permet posteriorment lligar-la amb adaptadors que tenen una "T", de forma similar a com es preparen les llibreries per truSeq. Els adaptadors són molècules de cadena senzilla de ADN que formen una forquilla intramolecular, que genera una molècula amb una forma característica coneguda com *SMRTbell DNA*.

La llibreria sintetitzada es col·loca sobre un xip de vidre que conté múltiples guies d'ona en mode zero (o *"zero-mode waveguides",* ZMW). Les ZMW són estructures circulars amb forma de pouet d'uns 70nm de diàmetre i 100nm de profunditat que permeten confinar els nanofotons en un volum aproximat de $20x10^{-21}$ L. En cadascun dels ZMW hi cap únicament una polimerasa (phi29) i una molècula de ADN motlle. Els pouets es tracten prèviament amb estreptovidina que permetrà immobilitzar el binomi polimerasa-SMRTbell DNA.

Finalitzada la fase de fixació la polimerasa incorpora els nucleòtids, específicament marcats cadascun amb un fluoròfor diferent, a la cadena replicada d'ADN. En el procés d'addició dels nucleòtids, aquests s'acosten a la base del pou ZMW on hi ha una càmera d'alta resolució que captura el canvi de fluorescència que emet el nucleòtid en alliberar un fòsfor (**Figura 21**). Aquest procés es repeteix de forma seqüencial per desxifrar la seqüència completa. La polimerasa utilitzada té una alta processivitat i, si el tamany d'insert és petit, permet llegir-lo tant en sentit forward com reverse múltiples vegades. Això permet obtenir una major precisió per base seqüenciada (Rodríguez-Ezpeleta et al., 2012).

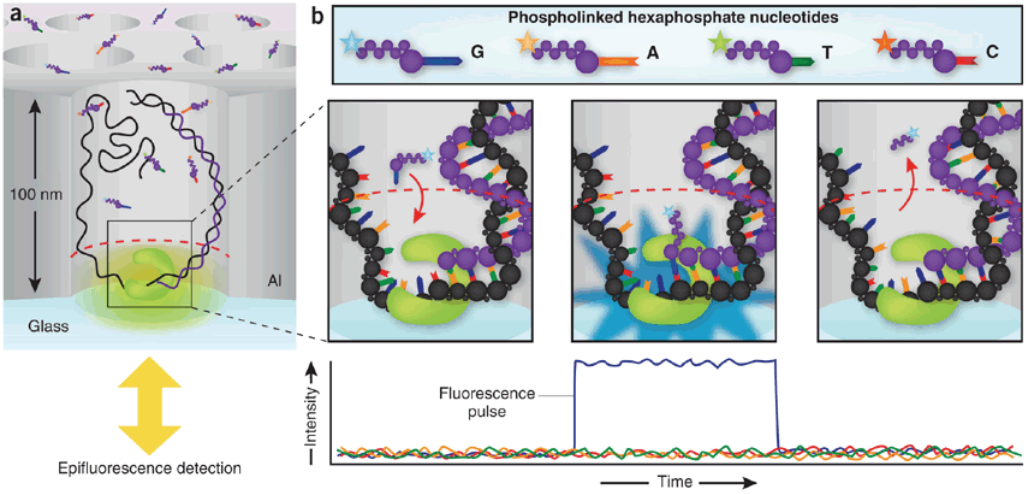


Figura 21. Diagrama de representació del funcionament de la plataforma PacBio (Metzker, 2010)

2.3.2.4. Oxford Nanopore

La tecnologia Nanopore es basa en l'anàlisi i registre dels canvis de corrent elèctric provocats pel pas dels nucleòtids d'una molècula a través d'un nanopor. Un nanopor és un petit forat amb un diàmetre intern menor a 1 nanòmetre. Algunes proteïnes transmembrana tenen la propietat d'actuar com a nanopors. Els canvis de corrent generats permeten discriminar el tipus de nucleòtid que passa pel nanopor i al processar aquests impulsos es pot determinar les bases que composen la cadena de ADN que s'està llegint.

Aquesta tècnica permet seqüenciar molècules senceres de longitud important sense necessitat de preparar o amplificar les mostres. Al 2014 la taxa d'error del sistema era del 30% (Kilianski et al., 2015); mentre que a finals de 2016 Oxford Nanopore va indicar que la taxa d'error havia disminuït fins un 2-13% en funció de la mostra, aconseguint alhora un rendiment de fins a 10Gbp.

2.3.2.5. Ion Personal Genome Machine (Ion Torrent)

El sistema conté una sèrie de xips semiconductors capaços de detectar canvis de pH originats en alliberar-se un protó d'hidrogen dels nucleòtids que s'incorporen a una cadena d'ADN. Aquest sistema no requereix de nucleòtids marcats ni enzims específics. L'ADN fragmentat es marca amb adaptadors en els dos extrems i s'enganxa a unes esferes que contenen una seqüència ssDNA complementaria a l'adaptador. Aquestes esferes amb les seqüències diana es carreguen en un xip microconductor que està dividit en molts pous on cadascuna d'aquestes boles s'acomoden. La base del xip té un sensor capaç de detectar els canvis de pH que es produeixin durant l'addició de nucleòtids i l'alliberació de protons, assignant a cada posició un nucleòtid diferent segons el canvi de conductivitat mesurat (**Figura 22**).

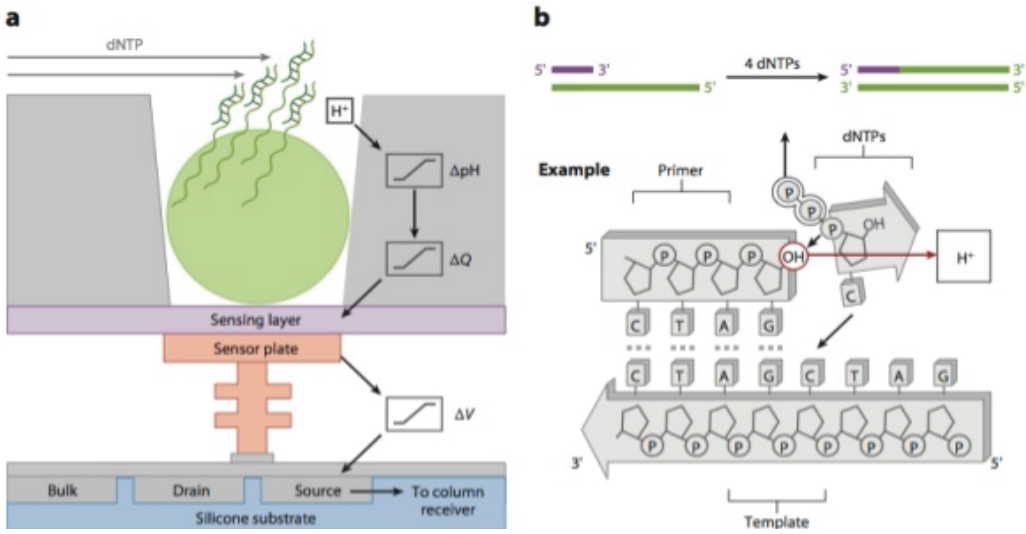


Figura 22. Representació de la plataforma ión Torrent (Mardis, 2013)

(a) Estructura del xip detector del Ion Torrent utilitzat en la lectura de bases per canvi de pH.

(b) Detecció del nucleotid per canvi de pH

Les principals tecnologies actualment disponibles per a la seqüenciació massiva es recullen a continuació a la Taula 5.

Taula 5. Principals tecnologies disponibles per a fer seqüenciació massiva.

Plataforma	Companyia	Llibreria	Suport	Generació	Reacció	Detecció
GS FLX o 454	Roche	Adaptadors lineals	Placa picotiter	PCR emulsió	Pirosequenciació	Nucleòtids marcats fluorescents
Mi-Seq/Hi-Seq	Illumina	Nextera/TruSeq	*Flow cell*	Bridge PCR	Síntesi	Nucleòtids terminadors marcats fluorescentment
PacBio RS	Pacific biociences	*Bubble adapters*	*Zero mode waveguide*	Single molecule	Síntesi	Nucleòtids lligats a fluoròfors
Ion torrent	Thermo Fisher	Adaptadors lineals acoplats a *beads*	Microxip semiconductor	PCR emulsió (*beads*)	Síntesi	Sensor iònic ISFET, detecció H^+
CGA platform	Complete genomics	Adaptadors Circulars	DNA nanoballs	RCA	Lligació	Sondes fluoròfores marcades
SOLiD	Applied Biosystems	Adaptadors lineals	*Flow Cell*	PCR emulsió	Lligació	Sondes fluoròfores marcades
Mini-Ion	Nanopore sequencing	No necessari	*Flow cell*	No amplificació	Pas a través del nanopor	Canvis corrent detectats pel nanopor
Oxford nanopore	Nanopore sequencing	No necessari	*Flow cell*	No amplificació	Pas a través del nanopor	Canvis corrent detectats pel nanopor

2.3.3. Anàlisi bioinformàtica de dades

Les noves tecnologies de seqüenciació generen una enorme quantitat de dades com ja s'ha mostrat en els apartats anteriors. Les seqüències generades s'anomenen *reads* i tenen longitud variable en funció de la plataforma utilitzada per seqüenciar i els protocols aplicats. Independentment de l'experiment de seqüenciació en sí mateix, les seqüències dels *reads* s'han de processar mitjançant l'ús de software específic. En primer lloc, s'han d'eliminar les seqüències romanents dels adaptadors. En segon lloc, s'ha d'avaluar la qualitat dels nucleòtids que composen la seqüència, per exemple fent servir el sistema de puntuacions de qualitat *Phred*, que mesura la probabilitat d'error d'una base en una posició concreta a través d'una escala logarítmica (Ewing et al., 1998). Els *reads* seran processats en funció de la seva qualitat eliminant les zones de baixa qualitat però també de baixa complexitat; aquest darrer aspecte és important ja

que es redueix el nombre de seqüències repetitives que poden afectar a la resta d'anàlisis computacionals "*downstream*". Mitjançant l'ús d'ensambladors es reconstrueixen seqüències de major tamany anomenades *contigs*. El procés d'ensamblat pot veure's dificultat per la curta longitud dels segments, els errors de seqüenciació, estructures repetitives dins del genoma, el volum de dades generades, i/o la variabilitat natural de la pròpia seqüència quan la mostra conté una barreja d'individus de la mateixa espècie (per exemple en el cas de tenir quasiespècies) o d'espècies diferents (ja que podem tenir dominis funcionals molt conservats o molt divergents amb famílies gèniques que evolucionen a diferents ritmes). Tots aquests problemes han afectat de forma important als estudis de metagenòmica, on milions d'organismes diferents es poden trobar presents en una simple mostra. Per superar aquestes dificultats s'han creat programes d'ensamblat específics per a treballar amb metagenomes, com MetaVelvet (Afiahayati et al., 2015; Namiki et al., 2012), MetaSpades (Nurk et al., 2017), i Ray-Meta Ray (Boisvert et al., 2012).

Els ensambladors es basen en tres estratègies algorítmiques, o abordatges, diferents (Nagarajan and Pop, 2013), els quals es mencionem a continuació:

Greedy. L'ensamblador conecta aquells *reads* que presenten un millor solapament de forma iterativa, mentre no contradiguin l'ensamblat ja construït prèviament. És l'algorisme més senzill i intuitiu. Són tècniques molt utilitzades en *reads* obtinguts pel mètode Sanger però poc eficients per ensamblar genomes grans a partir de milions de *reads* molt curts (com els que genera Illumina).

OLC (Overlap-Layout-Concensus). El mètode identifica aquells *reads* que se solapen suficientment bé i organitza la informació en un *graf* (una estructura de dades que modela matemàticament una xarxa) que relaciona els nodes entre ells a través d'un connector (*Edge*). L'algorisme aplicat traça el millor camí que permet enllaçar tots els nodes entre sí; aquest recorregut és el que permet generar el *contig* a partir d'un conjunt de *reads* solapants. Aquest sistema té en

consideració la relació global entre els *reads* a diferencia dels ensambladors basats en algorismes *Greedy*.

Grafs "De Bruijn". Algorismes molt complexes en el que es basen la majoria d'ensambladors moderns que treballen amb dades de seqüenciació massiva. Els *reads* es descomponen en fragments més curts de longitud *k* anomenats "*k-mers"*. De manera similar al mètode OLC, els nodes en el gràfic representen *k-mers* en lloc de seqüències completes de *reads*, i els connectors indiquen quins *k-mers* adjacents es solapen per *k-1* lletres en lloc de representar solapaments de *reads*. En aquest cas la longitud del *k-mer* correlaciona amb la longitud del solapament que l'ensamblador és capaç de detectar, que és k-1. En aquesta metodologia no es modelen directament els *reads*, sinó que aquests són representats pels connectors del diagrama o graf de Bruijn.

En un pas posterior a l'ensamblat, els *contigs* generats poden ser ordenats i enllaçats entre ells per generar seqüències més grans, conegudes com *scaffolds.* Finalment, les seqüències processades que han passat tots els criteris de qualitat s'han d'assignar taxonòmicament, mitjançant programes d'alineament especials, també coneguts com a programes de mapat, com Bowtie2 (Langmead et al., 2009) o per similitud de nucleòtids o proteïnes, fent servir BLASTN o BLASTX (Altschul et al., 1997, 1990). Aquesta cerca es realitza comparant les seqüències obtingudes pel protocol d'ensamblat contra bases de dades de seqüències conegudes, com pot ser NCBI-GenBank o UniProt, o la nostra pròpia base de dades local on s'hagin filtrat seqüències específiques de genomes, gens i/o proteïnes virals. Les seqüències amb una menor similitud ens poden permetre detectar nous membres de famílies, mentre que aquelles seqüències que queden sense anotar podrien representar noves espècies i podrien ser candidates a posteriors validacions experimentals per caracteritzar els organismes en qüestió, tasca molt més complexa i laboriosa.

La present Figura 23 mostra els possibles passos a seguir després de fer un *run* amb Illumina Mi-seq, representant un exemple de pipeline bioinformàtica aplicada a la classificació de seqüències víriques.

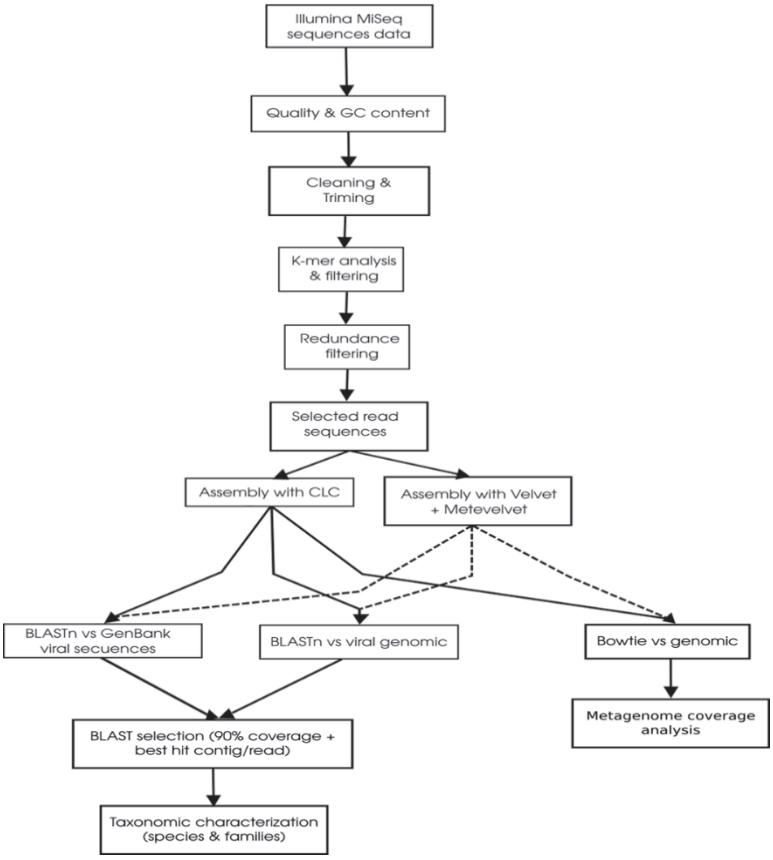


Figura 23. Diagrama de flux del protocol bioinformàtic a plicat durant la present tesi doctoral

OBJECTIUS

OBJECTIUS

L'objectiu global plantejat en aquesta tesi ha estat l'anàlisi dels patògens virals que a través de l'aigua residual poden contaminar aigua de reg i aliments.

Els objectius específics són:

1) Estudiar l'eficiència en la reducció de virus i altres patògens en un sistema de llacunatge per producció d'aigua regenerada.

2) Avaluar l'efecte que tenen sobre les poblacions virals diferents mètodes de concentració i extracció de virus quan s'apliquen tècniques de seqüenciació massiva.

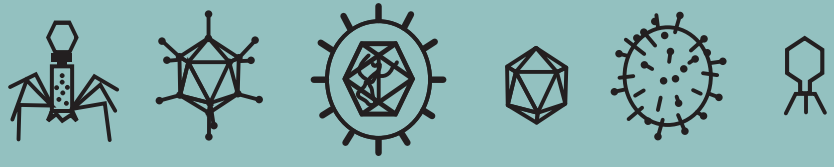
3) Desenvolupar metodologies eficients per l'estudi del viroma excretat en aigües residuals.

4) Descriure el viroma de l'aigua residual urbana avaluant diferències entre estacions i caracteritzant virus nous i emergents.

5) Caracteritzar la contribució de l'orina al viroma de l'aigua residual urbana.

6) Anàlisi metagenòmica de la contaminació viral en vegetals frescos experimentalment cultivats i regats amb aigua de riu, aplicacions de la metagenòmica a la seguretat alimentària.

INFORME DELS ARTICLES



INFORME DELS ARTICLES

1.1. Llista d'articles inclosos a la tesi

La present tesi està fomentada en les següents publicacions detallades per ordre de presentació:

- Fernandez-Cassi, X., Silvera, C., Cervero-Aragó, S., Rusiñol, M., Latif-Eugeni, F., Bruguera-Casamada, C., Civit, S., Araujo, R.M., Figueras, M.J., Girones, R., Bofill-Mas, S., 2016. Evaluation of the microbiological quality of reclaimed water produced from a lagooning system. Environ. Sci. Pollut. Res. 1–18. doi:10.1007/s11356-016-6812-0

- Hjelmsø, M.H., Hellmér, M., Fernandez-Cassi, X., Timoneda, N., Lukjancenko, O., Seidel, M., Elsässer, D., Aarestrup, F.M., Löfström, C., Bofill-Mas, S., Abril, J.F., Girones, R., Schultz, A.C., 2017. Evaluation of Methods for the Concentration and Extraction of Viruses from Sewage in the Context of Metagenomic Sequencing. PLoS One 12, e0170199. doi:10.1371/journal.pone.0170199

- Fernandez-Cassi, X., Timoneda, N., Martínez-Puchol, S., Rusiñol, M., Rodríguez-Manzano, J., Figuerola, N., Purcell, R.H., Abril, J.F., Bofill-Mas, S., Girones, R. Metagenomics for the study of the viral contamination in urban wastewater. Manuscrit sotmès al 2017.

- 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. Manuscrit sotmès al 2017.

1.2. Informe de coautoria

Fernandez-Cassi, X., Silvera, C., Cervero-Aragó, S., Rusiñol, M., Latif-Eugeni, F., Bruguera-Casamada, C., Civit, S., Araujo, R.M., Figueras, M.J., Girones, R., Bofill-Mas, S., 2016. Evaluation of the microbiological quality of reclaimed water produced from a lagooning system. Environ. Sci. Pollut. Res. 1–18. doi:10.1007/s11356-016-6812-0

El present article s'emmarca dintre el projecte del ministeri d'Economia i Competitivitat (AGL2011-30461-CO2-01). El doctorand ha participat activament en les tasques de mostreig, processament i anàlisi molecular de virus de les mostres. També s'ha encarregat dels assajos d'infectivitat i de part de l'anàlisi estadística. Addicionalment, ha escrit el manuscrit de l'article sota la supervisió i ajuda de la resta de co-autors.

Hjelmsø, M.H., Hellmér, M., Fernandez-Cassi, X., Timoneda, N., Lukjancenko, O., Seidel, M., Elsässer, D., Aarestrup, F.M., Löfström, C., Bofill-Mas, S., Abril, J.F., Girones, R., Schultz, A.C., 2017. Evaluation of Methods for the Concentration and Extraction of Viruses from Sewage in the Context of Metagenomic Sequencing. PLoS One 12, e0170199. doi:10.1371/journal.pone.0170199

El present article s'emmarca dintre el projecte europeu METAWATER en el que el doctorand ha treballat activament encarregant-se de l'elaboració dels protocols de seqüenciació massiva per a l'estudi de virus. En concret, el doctorand ha participat activament en el disseny experimental amb els dos coautors principals. També ha participat de forma activa en l'anàlisi de dades i redacció del manuscrit.

Fernandez-Cassi, X., Timoneda, N., Martínez-Puchol, S., Rusiñol, M., Rodríguez-Manzano, J., Figuerola, N., Purcell, R.H., Abril, J.F., Bofill-Mas, S., Girones, R. Metagenomics for the study of the viral contamination in urban wastewater. Manuscrit sotmès al 2017.

El present article s'emmarca dintre el projecte del ministeri d'Economia i Competitivitat (AGL2011-30461-CO2-01). El doctorand s'ha encarregat de la gestió i recollida de mostres així com també de la part experimental que fonamenta l'article. S'ha encarregat de l'anàlisi informàtica de les seqüències conjuntament amb el segon autor. Finalment, ha elaborat el manuscrit sota la supervisió de les seves codirectores de tesi.

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. Manuscrit sotmès al 2017.

El present article s'emmarca dintre el projecte Recercaixa. El doctorand s'ha encarregat de la part experimental i dels mostrejos en els que es sustenta l'article. S'ha encarregat de l'anàlisi informàtica de les seqüències conjuntament amb el segon autor. Finalment, ha elaborat el manuscrit sota la supervisió de les seves codirectores de tesi.

Signat,

Dra. Rosina Girones Llop

Barcelona, Maig 2017

Dra. Sílvia Bofill Mas

Barcelona, Maig 2017

1.3. Informe sobre el factor d'impacte

Els treballs que formen part de la present tesi doctoral s'han publicat o s'han sotmès a publicació en revistes científiques rellevants per a la línia de recerca en la qual el doctorand ha participat durant els anys de tesi.

L'article **"Evaluation of the microbiological quality of reclaimed water produced from a lagooning System"** s'ha publicat a la revista *Environmental and Pollution Research* l'any 2016. L'index d'impacte d'aquesta revista a l'any 2015 era de **2,760**.

L'article **"Evaluation of Methods for the Concentration and Extraction of Viruses from Sewage in the Context of Metagenomic Sequencing"** s'ha publicat a la revista PLOSone l'any 2017. L'index d'impacte d'aquesta revista a l'any 2015 era de **3,057**.

L'article **"Metagenomics for the study of the viral contamination in urban wastewater"** s'ha sotmès a la revista *Science of the Total Environment.* L'index d'impacte al 2015 d'aquesta revista era **3,976.**

L'article **"A metagenomic assessment of viral contamination on fresh parsley plants irrigated with fecally tainted river wàter"** s'ha sotmès a la revista *International Journal of Food Microbiology* l'any 2017. L'index d'impacte d'aquesta revista a l'any 2015 era de **3,445**.

Signat,

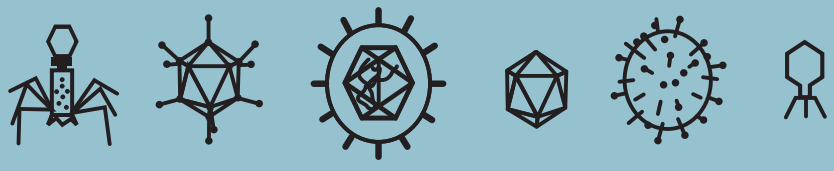
Dra. Rosina Girones Llop

Barcelona, Maig 2017

Dra. Sílvia Bofill Mas

Barcelona, Maig 2017

ARTICLES



ARTICLE I

Avaluació de la qualitat micobiològica de l'aigua regenerada produïda en un sistema de llacunatge

ARTICLES

1. Avaluació de la qualitat microbiològica de l'aigua regenerada produïda en un sistema de llacunatge.

Les previsions del IPCC referents al consum d'aigua de cara al 2025 assenyalen un increment en el consum, principalment associat a un creixement de la població mundial. A la vegada, els períodes de sequera associats al canvi climàtic compliquen la futura disponibilitat d'aigua. Aquest escenari reforça la idea que és necessària una major reducció, reutilització i reciclatge dels recursos hídrics disponibles. L'ús d'aigua reciclada o regenerada, obtinguda a partir d'aigua residual, representa una oportunitat única per combatre la manca d'aigua. No obstant, els tractaments de producció d'aigua regenerada com la cloració o l'ús de llum ultraviolada són econòmicament i energèticament costos. En aquest sentit l'aplicació de sistemes de tractament com el llacunatge, amb costos econòmics i mediambientals menors, representen una alternativa interessant per a produir aigua regenerada. Aquesta font d'aigua es regula sota el Reial Decret 1620/2007 que n'estableix els usos en funció de la seva càrrega microbiològica, expressada a través dels indicadors bacterians tradicionals com *E.coli* o enterococs intestinals. Malgrat l'existència d'aquests paràmetres, l'absència d'aquests indicadors no té perquè correlacionar amb la d'altres patògens bacterians, vírics o protozous, alguns dels quals són capaços de recréixer sota determinades condicions ambientals o tenen una major resistència en l'ambient.

Amb la intenció d'esbrinar la qualitat microbiològica de l'aigua regenerada produïda en el sistema de llacunatge es va avaluar la presència de microorganismes, alguns d'ells patògens, a l'entrada i a la sortida del sistema de llacunes de forma mensual durant un any. Entre els virus estudiats trobem els Adenovirus humans, els poliomavirus JC, el virus de l'hepatitis E i els norovirus

humans dels genogrups I i II. També s'ha estudiat la presència d'heteròtrofs totals, amebes de vida lliure i bacteris dels gènere *Legionella, Arcobacter i Aeromonas,* i dels FIB *E.coli* i Enteroccocs intestinals.

A partir de l'anàlisi de mostres d'aigua a l'entrada i la sortida del sistema de llacunatge s'observa que s'aconsegueix reduir de forma estadísticament significativa en 1,18 i 0,64 logaritmes el nombre de còpies genòmiques d'HAdV i JCPyV, respectivament. Malgrat la baixa reducció obtinguda a través de les proves moleculars, els assajos d'infectivitat aplicats en mostres d'efluent de la llacuna no han permès detectar HAdV infecciosos. El sistema de llacunatge assoleix reduccions logarítmiques més elevades en indicadors bacterians tradicionals, obtenint 2,58 i 1,65 logaritmes de reducció per *E.coli* i Enterococs intestinals, respectivament. No obstant, més de la meitat de les mostres d'efluent de la llacuna contenen *E.coli* en uns nivells superiors als estipulats en el RD1620/2007 per al reg de vegetals consumits crus. Els resultats obtinguts senyalen la necessitat de dur a terme més estudis que ajudin a entendre els mecanismes d'eliminació que intervenen en els sistemes de llacunatge per a poder optimitzar aquests sistemes i produir una aigua regenerada de qualitat microbiològica suficient per al seu ús com aigua de reg.

Environ Sci Pollut Res (2016) 23:16816–16833
DOI 10.1007/s11356-016-6812-0

RESEARCH ARTICLE

Evaluation of the microbiological quality of reclaimed water produced from a lagooning system

X. Fernandez-Cassi[1] · C. Silvera[2] · S. Cervero-Aragó[1,3,4] · M. Rusiñol[1] · F. Latif-Eugeni[2] · C. Bruguera-Casamada[1] · S. Civit[5] · R. M. Araujo[1] · M. J. Figueras[2] · R. Girones[1] · S. Bofill-Mas[1]

Received: 7 January 2016 / Accepted: 2 May 2016 / Published online: 19 May 2016

Abstract The use of lagooning as a complementary natural method of treating secondary effluents of wastewater treatment plants has been employed as an affordable and easy means of producing reclaimed water. However, using reclaimed water for some purposes, for example, for food irrigation, presents some risks if the effluents contain microbial pathogens. Classical bacterial indicators that are used to assess faecal contamination in water do not always properly indicate the presence of bacterial or viral pathogens. In the current study, the presence of faecal indicator bacteria (FIB), heterotrophic bacterial counts (HBC), pathogens and opportunistic pathogens, such as *Legionella* spp., *Aeromonas* spp., *Arcobacter* spp., free-living amoeba (FLA), several viral indicators (human adenovirus and polyomavirus JC) and viral pathogens (noroviruses and hepatitis E virus) were analysed for 1 year in inlet and outlet water to assess the removal efficiency of a lagooning system. We observed 2.58 (1.17–4.59) and 1.65 (0.15–3.14) log reductions in *Escherichia coli* (EC) and intestinal enterococci (IE), respectively, between the inlet and outlet samples. Genomic copies of the viruses were log reduced by 1.18 (0.24–2.93), 0.64 (0.12–1.97), 0.45 (0.04–2.54) and 0.72 (0.22–2.50) for human adenovirus (HAdV), JC polyomavirus (JCPyV) and human noroviruses (NoV GI and GII), respectively. No regrowth of opportunistic pathogens was observed within the system. FLA, detected in all samples, did not show a clear trend. The reduction of faecal pathogens was irregular with 6 out of 12 samples and 4 out of 12 samples exceeding the EC and IE values, specified in the Spanish legislation for reclaimed water (RD 1620/2007). This data evidences that there is a need for more studies to evaluate the removal mechanisms of lagooning systems in order to optimize pathogen reduction. Moreover, surveillance of water used to irrigate raw edible vegetables should be conducted to ensure the fulfilment of the microbial requirements for the production of safe reclaimed water.

Keywords Lagooning · Bacterial removal · Virus removal · Regrowth · Reclaimed water · Faecal indicator bacteria · Human adenovirus

Responsible editor: Robert Duran

Electronic supplementary material The online version of this article (doi:10.1007/s11356-016-6812-0) contains supplementary material, which is available to authorized users.

✉ X. Fernandez-Cassi
xaviako@gmail.com

[1] Department of Microbiology, Faculty of Biology, University of Barcelona, Barcelona, Spain

[2] Unitat de Microbiologia, Departament de Ciències Médiques Bàsiques, Facultat de Medicina i Ciències de la Salut, IISPV, Universitat Rovira i Virgili, Reus, Spain

[3] Water Hygiene, Institute for Hygiene and Applied Immunology, Medical University of Vienna, Vienna, Austria

[4] ICC Water and Health, Vienna, Austria

[5] Department of Statistics, Faculty of Biology, University of Barcelona, Barcelona, Spain

Introduction

Water scarcity is a major problem worldwide, with an estimated 2.7 million people living close to river basins that are affected by severe water need (Oki and Kanae 2006; Hoekstra et al. 2012). The United Nations estimates that the world population will reach 9 billion in approximately 2050 and that water availability will therefore decrease. Current projections indicate that there will be a 55 % increase in water demand between 2000 and 2050 (Gurría 2012). It is clear that reducing

water use, recycling and reusing water resources are and will continue to be a priority in the near future.

The use of recycled treated wastewater, also called reclaimed water, for irrigation has proven to be a good strategy for reducing water scarcity. Reclaimed water is a source of water that is independent of weather conditions and includes additional nutritional input that can produce better vegetables, field crops, and fruits (Wheaton et al. 2001; Dare 2015). However, the use of improperly treated reclaimed water for food irrigation presents a risk to human health if the pathogenic microorganisms contained in the water are not inactivated (Carter 2005; Riera-Montes et al. 2011). The increase of fresh food consumption has been linked to an increase in foodborne outbreaks (Sivapalasingam et al. 2004; Kozak et al. 2013; Callejón et al. 2015). One of the sources of food viral contamination is irrigation water. Maunula et al. (2013) found that HAdV was present in 9.5 % of irrigation water used to irrigate berries proving that it was faecally contaminated. Recently, a major outbreak of norovirus-related gastroenteritis affected nearly 11,000 people in Germany and was linked to strawberries that were probably irrigated with contaminated water during fruit production (Bernard et al. 2014). Although it is not stated in the report if the water used for irrigation was reclaimed or not, this incident provides a good example of how contaminated water or improperly treated reclaimed water can result in a health risk.

The Spanish legislation on water reuse (Royal Decree 1620/2007) has represented an important advancement to standardize reuse practices differentiating 14 uses under five main areas: urban, agricultural irrigation, industrial, recreational and environmental. To produce reclaimed water with a higher quality and to minimize/prevent the risk of human infections, several processes, such as microfiltration and ultrafiltration, are being applied after the secondary treatment in wastewater treatment plants (WWTP). The technological needs and high costs of these treatments pose substantial challenges, especially in some countries, because of economic constraints. Among the different low-cost methods used to produce reclaimed water, storing treated or untreated wastewater in stabilization ponds, polishing ponds, or natural or artificial lagoons appears to be an effective solution to overcome these challenges (Campos et al. 2002; Oragui et al. 2011). These methods are affordable in developing countries and small communities, and they result in high microbial inactivation rates via the use of cheap technological approaches with low operational costs (Mara et al. 1992; Peña et al. 2000).

Traditionally, the efficiency of microbial removal for these methods has been evaluated using faecal indicator bacteria (FIB) such as faecal coliforms. However, the presence of faecal coliforms is not always correlated with the presence of viral pathogens, such as noroviruses (Gerba et al. 1979; Marzouk 1980; Pusch et al. 2005; Jiang 2006; Bofill-Mas et al. 2013) or other pathogenic bacteria, such as *Legionella* spp., *Aeromonas* spp., which are indigenous to freshwater ecosystems, and *Arcobacter* spp., which is considered to be an opportunistic pathogen and a signature indicator of sewage contamination (Harwood et al. 2005; Collado and Figueras 2011). The weakness of this correlation has been attributed to differences in the survival rates of these species and differences in the efficiencies of treatments used in WWTPs. This lack of correlation has also been described in water stabilization ponds and lagooning systems in previous studies (Mara and Pearson 1987; Donnison and Ross 1995). Nevertheless, new studies based on molecular techniques are needed to confirm or deny a lack of correlation between these biological indicators. For example, in a recent study published by Jurzik et al. (2015), the use of polishing ponds as a tertiary treatment resulted in a reduction of 1.84–2.65 log units of bacteria and bacteriophages without reducing/affecting the concentrations of animal viruses. However, in the later study, viruses were tested by molecular methods and infectivity data was not included. The information derived from molecular methods is useful, especially for those viral agents which are not cultivable, such as NoV or bacterial species that might enter into a viable but not cultivable (VBNC) state. However, this is a limiting factor when evaluating the removal efficiency of a specific water treatment process because non-infective viral genomic material can be detected by q(RT)-PCR methods. To overcome this limitation, cell culture methods were applied, for HAdV, to test the infectivity of viral concentrates. Human adenovirus is widely used as a human viral faecal marker showing a high occurrence during all periods of the year (Gerba et al. 1979; Lipp et al. 2001; Bofill-Mas et al. 2013). The presence of human adenovirus infectious viral particles is important when evaluating the risks derived from water reuse, for example, when testing water used for crops irrigation.

The current study describes the removal efficiency of a lagooning system by analysing the presence of faecal viral markers, including human adenoviruses (HAdVs) and JC polyomavirus (JCPyV), heterotrophic bacteria counts (HBC), and classical FIBs, (*E. coli* (EC) and intestinal enterococci (IE)). In addition, the presence of pathogenic noroviruses (GI and GII), the hepatitis E virus (HEV), potential bacterial pathogens, such as *Arcobacter* spp., *Aeromonas spp*., *Legionella* spp. and free-living amoeba were analysed to determine whether these water systems are reservoirs or niches that might promote the regrowth of pathogenic microorganisms and thereby represent a new threat with regard to further water reuse.

Materials and methods

Description of the site and sampling program

The lagooning system evaluated is situated in south Catalonia (Spain) in a zone with typical Mediterranean weather. The lagooning system has a theoretical total volume capacity of 24,087 m^3 and a surface area of 16,864 m^2. The system receives water from the secondary outlet of a WWTP that treats a volume of 25,000 m^3/day of raw sewage from approximately 200,000 inhabitants. The flow rates were measured in continuous using a magnetic flow meter. The volume of water entering the lagoons is registered daily and was provided by the WWTP, and the data is presented in Table 1. Once the water has been treated with a conventional secondary treatment (activated sludge), its entry into the stabilization ponds of the lagooning tertiary treatment depends on water demands and is regulated by a water level control system. The system is composed of four lagoons with depths ranging from 1.95 to 3.15 m. This volume of water, which does not take evaporation into account, provides an indirect measure of reclaimed water produced and supplied to 140 users who use this water as the main irrigation source for their olive and hazelnut trees and vineyards, which cover an area of 135 ha.

One year of sampling, from September 2012 to August 2013, was completed. Two-litre water samples were collected monthly from the secondary outlet of the WWTP as it entered the lagooning system (lagooning inlet), and 2 L was collected from the tertiary effluent after it exited the stabilization ponds (lagooning outlet). A volume of 1 L was obtained from each sampling point and used for the viral analysis, 500 mL were used to analyse the samples for the presence of FIB, *Aeromonas* spp. and *Arcobacter* spp., and 500 mL were used to analyse the samples for the presence of HBC, *Legionella* spp. and free-living amoebas. Water temperature, pH and conductivity were measured using the corresponding probes (XS instruments device) and following the US EPA guidelines 150.1 and 120.1, respectively. Turbidity was determined by nephelometry using a hanna instrument and following the US EPA 180.1 guideline. All the variables were recorded during sampling. Precipitation and solar radiation data were collected from the Catalan Meteorological Institute (http://www.meteo.cat/servmet/index.html). All of this information is presented in Table 1.

Viral analysis

Viral concentration and nucleic acid extraction

The viruses present in 1 L samples were concentrated using skimmed milk organic flocculation. The method has a recovery efficiency of about 50 % (20–95 %) (Calgua et al. 2008, 2013). All samples were adjusted to a conductivity of 1.5 mS/cm^2 and acidified to a pH 3.5 using 1 N HCl. Briefly, a suspension of skimmed milk was prepared by adding 10 g of skimmed milk powder (Difco, Detroit, MI, USA) to 1 L of

Table 1 Physicochemical parameters and environmental factors analysed during the study period

Sampling date	Water matrices	27/09/12	29/10/12	27/11/12	17/12/12	28/01/13	25/02/13	18/03/13	29/04/13	20/05/13	17/06/13	29/07/13	26/08/13
pH	LI[a]	7.34	7.31	7.55	7.15	7.72	7.33	7.53	7.39	7.66	7.66	7.44	7.3
	LO[b]	7.56	8.28	7.20	7.57	8.16	7.92	7.88	7.88	8.28	8.27	7.66	7.67
Conductivity (mS/cm^2)	LI[a]	1313	1620	1350	1585	1520	1208	950	909	1214	1618	1552	1567
	LO[b]	1476	1455	1079	1540	1492	1451	1236	1576	1494	1574	1715	1621
Turbidity (NTU)	LI[a]	16.8	5.15	5.49	8.74	13.0	15.7	6.23	9.0	6.8	14.13	7.39	5.66
	LO[b]	8.34	19.4	13.5	9.9	12.4	29.5	10.5	10.0	6.78	12.7	19.3	15.7
Water temperature (°C)	LI[a]	23.0	14.6	12	9.9	15.0	15.5	18.0	17.0	21.0	26.0	27.0	26.0
	LO[b]	25.4	18.8	10	8.4	10.0	10.0	13.5	18.0	20.0	28.0	28.0	26.0
Average atmospheric temperature 72 h before sampling (°C)	LI[a]	20.57	13.10	12.77	12.93	10.27	5.87	9.30	11.27	13.60	21.83	26.43	22.87
	LO[b]	20.57	13.10	12.77	12.93	10.27	5.87	9.30	11.27	13.60	21.83	26.43	22.87
Accumulated precipitations 72 h before sampling (mm)	LI[a]	0	0	0	0.9	0	7	0	8.6	1.8	0.4	0	0
	LO[b]	0	0	0	0.9	0	7	0	8.6	1.8	0.4	0	0
Accumulated solar radiation 72 h before sampling (MJ/m^2)	LI[a]	20.6	7.6	7	2.9	12	2.1	21	8.8	27.3	29.6	25.3	17.4
	LO[b]	20.6	7.6	7	2.9	12	2.1	21	8.8	27.3	29.6	25.3	17.4
Total monthly consumption (m^3)		58478	21100	18720	12017	19610	20159	10960	26381	21600	47040	86989	75937
Theoretical hydraulic retention time (months)		0.85	2.37	2.67	2.94	2.54	2.48	4.56	1.89	2.31	1.06	0.57	0.66

[a] Lagooning inlet

[b] Lagooning outlet

artificial seawater (Sigma-Aldrich Chemie GMBH, Steinheim, Germany), and the solution was then adjusted to pH 3.5 using 1 N HCl to obtain a pre-flocculated 1 % (*w*/*v*) skimmed milk solution (PSM). Then, 10 mL of PSM was added to previously conditioned samples to obtain a final concentration of 0.01 % of skimmed milk. Samples were kept for 8 h while stirring at room temperature, and flocks were allowed to sediment by gravity during 8 h. The supernatant was carefully removed, and the remaining 500 mL of the solution were centrifuged at 8000×*g* for 30 min at 4 °C. Pellets were suspended in 1 mL of phosphate buffer (pH 7.5) and stored at −80 °C until nucleic acid (NA) extractions were performed. A negative concentration control for each process sample was also included. For these, we used tap water as the matrix, and we first neutralized the free chlorine by adding 100 mL of a 10 % sodium thiosulfate solution.

To extract viral nucleic acids, 140 μL of viral concentrates were processed using a QIAamp® Viral RNA Mini Kit (Qiagen, Valencia, CA, USA) and the automated QIACube system (Qiagen, Valencia, CA, USA) following the manufacturer's instructions. Nucleic acids were stored at −80 °C until analysed. A negative control consisting of DNAse/RNAse-free molecular water was included in each extraction batch.

Quantitative and nested PCR assays to evaluate viruses

Samples were analysed to determine the presence and concentrations of viral faecal markers and other pathogenic viruses.

Specific real-time qPCR assays were used to quantify the viral faecal markers HAdV and JCPyV (Bofill-Mas et al. 2006; Hernroth et al. 2002; Pal et al. 2006) using TaqMan® Environmental Master Mix 2.0 (Life technologies, Foster City, CA, USA). Real-time primers and probes for HAdV can amplify A, C, D, E, F and some B HAdV serotypes, so the most frequently described HAdV can be detected with this assay. Specific RT-qPCR assays were performed to quantify the levels of the human norovirus genogroups I and II (NoVGI and NoVGII) (Kageyama et al. 2003; Loisy et al. 2005) and the RNA UltraSense™ One-Step Quantitative RT-PCR System (Invitrogen, Carlsbad, CA, USA). All samples were analysed in duplicate using undiluted and 1:10 dilutions of the nucleic acids extracts. Dilutions were tested to detect the potential inhibition of amplification resulting from the presence of inhibiting substances in the undiluted samples. The analysis of direct and ten-fold dilutions of environmental samples may indicate the presence of enzymatic inhibition if a difference of Cq between direct and diluted sample is lower than 3.3. In this case, the value considered was the one obtained by testing the diluted sample and more dilutions of that sample were tested to rule out if inhibition was still occurring. Also, external amplification controls were used to evaluate the potential inhibitory capability of the studied samples by adding known amounts of standard plasmid (1×10^3 GC/reaction). A MX3000Pro qPCR sequence detector system (Stratagene, La Jolla, CA, USA) was used to quantify the levels of viral genomes in the samples.

Nested RT-PCR (nRT-PCR) assays were used to test for the presence of the hepatitis E virus (HEV), as previously described by Erker et al. (1999). Nested PCR was selected because low numbers of HEV were expected, and nested PCR facilitated sequencing analysis of the virus. Reverse transcription of the extracted RNA was performed using a one-step RT-PCR Kit (QIAGEN, Valencia, CA, USA), and semi-nested PCR was performed using AmpliTaq™ Gold DNA polymerase.

The limit of detection (LOD) in 100 mL of water of the (RT)qPCR assays that were used in this study was found to be 29 GC for HAdV, 80 GC for JCPyV, 343 GC for NoVGI and 229 GC for NoVGII following the FSA 2006 guidelines.

Controls for (RT)-qPCR assays

Plasmid DNA suspensions were used as positive controls and quantitative standards. For HAdV and JCPyV, the hexon region (8961 bp) of HAdV41 and the whole genome (5130 bp) of JCPyV Mad1 were cloned into the plasmid pBR322. The capsid protein regions of NoVGI.4 (2931 bp) and NoVGII.13 (3004 bp) were cloned into the pTrueBlue®-Pvu II vector and used as the qRT-PCR standard.

To reduce the possibility of DNA contamination in the laboratory, 10 μg of each plasmid DNA was linearized using specific restriction enzymes as follows: BamHI for the HAdV41 plasmid (Promega, Madison, WI), NruI for the JCPyV plasmid (Promega, Madison, WI), SacI for the NoVGI plasmid and XhoI for the NoVGII plasmid (Promega, Madison, WI). The reaction products were subsequently purified and quantified. Samples and standard plasmids were added in two different rooms to avoid the possibility of contamination. Serial dilutions in TE buffer were performed using linearized standards ranging from 10^0 to 10^5 molecules per 5 or 10 μL (for viral RNA or DNA, respectively). Aliquot standard dilutions were stored in individual tubes at −80 °C until use.

All qPCR assays included non-template controls (NTC), and control extractions were included to evaluate any possible contamination during the extraction and amplification process. Moreover, all qPCR, RT-qPCR, nPCR and RT-PCR mixes, sample inoculations and standard additions were performed in separated areas to avoid any potential contamination. Negative PCR controls were also included for each analysis.

Infectivity assays in HAdV using the ICC-qPCR approach

An infectivity assay was performed for HAdV using the human embryonic kidney cell line *HEK 293A* (Life technologies, R705-07). Cells were infected with the four inlet samples that had the highest number of HAdV viral genomic copies and the four corresponding outlet samples. Cells were used from passages 12 to 15 and cultured using Dulbecco's

modified Eagle's medium (DMEM) containing a high concentration of glutamine (Glutamax, Life Technologies). The medium was supplemented with 10 % fetal bovine serum (Life Technologies), 1 % streptomycin-ampicillin and 1 % non-essential amino acids (Life Technologies), as previously described in the literature (Ogorzaly et al. 2013).

HEK 293A cells were seeded in 25 cm^2 cell culture flasks at a density of 5×10^4 cells/mL and incubated in 5 % CO_2 at 37 °C until confluence was achieved. Each environmental sample was analysed in two infected cell flasks (T0-1 h incubation and T8-8 day incubation). Cell culture flasks were infected using 100 μL of viral concentrate that was diluted in DMEM (1:1) to achieve a final infective solution of 200 μL because previous assays showed toxicity when they were infected with undiluted viral concentrates. Cells were incubated at 37 °C for 60 min. The cells were subsequently washed with PBS 1× three times to remove non-infective viral particles that were attached to cell surfaces. Finally, 5 mL of DMEM supplemented with 1 % non-essential amino acids, 2 % fetal bovine serum, 2 % streptomycin-ampicillin and 2 % kanamycin were added to the cell flasks. A negative (DMEM) and a positive control (HAdV35) were performed in parallel.

Positive samples were quantified with a most probable number (MPN) approach. Briefly, nine 25-cm^2 cell culture flasks were inoculated using a ten-fold dilution series (direct to 10^{-2}) in triplicate. All of the infected cell-cultured flasks were scratched after 3 days of incubation and analysed using 140 μL of the scratched cell suspension in a QIAamp Viral RNA mini Kit (QIAGEN). Negative samples had <8 MPN infective HAdV in 100 mL.

Cell cultures presumptively positive for adenovirus were analysed using a nested PCR (Allard et al. 2001), and the amplicon was sequenced to typify the most abundant HAdV types grown in cell cultures. The pair of primers selected were broad primers to amplify all HAdV known serotypes. The amplicons obtained after the nested-PCR were purified using a QIAquick PCR purification kit (QIAGEN, Inc.). The purified DNA was directly sequenced using an ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction kit version 3.1 with AmpliTaq® DNA polymerase FS (Applied Biosystems) according to the manufacturer's instructions. Conditions for the 25-cycle sequencing amplification were denaturing at 96 °C for 10 s, annealing for 5 s at 50 °C and extension at 60 °C for 4 min. Nested primers were used for the sequencing reactions at a concentration of 0.05 μM.

Sequencing results were checked using an ABI PRISM 377 automated sequencer (PerkinElmer, Applied Biosystems). Sequences were compared to the GenBank and European Molecular Biology Library (EMBL) using the basic National Center for Biotechnology Information (NCBI) BLAST programme (http://www.ncbi.nlm.nih.gov/BLAST/).

Bacteria and protozoa analysis

Heterotrophic bacteria quantification

Heterotrophic bacterial counts (HBC) were determined in accordance with International Organization for Standardization (1999) following the standards for water quality. Briefly, ten-fold dilution series were prepared in Ringer 1/4 (Scharlau Chemie; Barcelona, Spain), plated in plate count modified agar media (Scharlau Chemie; Barcelona, Spain) and incubated at 22 °C for 72 h as previously described (Serrano-Suárez et al. 2013) .

FIB

The determination of *E. coli* (EC) and intestinal enterococci (IE) was performed using 96-well microplate MPNs according to the International Organization for Standarization (1998a, 2012) methods (Bio-Rad, France), respectively. The detection method used for EC was based on the expression of the β-D-glucuronidase enzyme, while the expression of β-glucosidase was the target used to detect IE.

Detection and quantification of Arcobacter

The quantification of *Arcobacter* in reclaimed water was performed according to the MPN method using five replicate tubes. Each contained 2.5 mL of *Arcobacter* broth supplemented with CAT (cefoperazone, amphotericin B and teicoplanin) and 0.5 mL of a ten-fold serial dilution of the water samples. The tubes were incubated at 30 °C for 48 h. For tubes that showed turbidity, 200 μL of the broth was inoculated using passive filtration (0.45-μm membrane) on blood agar plates (Trypticase Soy Agar (TSA), BD, Spain) supplemented with 5 % sheep's blood. Eight presumptive *Arcobacter* colonies were selected for molecular identification from each positive sample. The 16S rDNA-RFLP method (Figueras et al. 2012) was used. The MPN values were calculated using MPN Build 23 software (Mike Curiale software), and the results were expressed as MPN/100 mL.

Detection and quantification of Aeromonas

All water samples were investigated for the presence of *Aeromonas* ($n = 24$), but quantification was performed according to the MPN method using five replicate tubes only between March 2013 and August 2013 ($n = 12$). Buffered peptone water (BPW, Oxoid, UK) was used to prepare ten-fold serial dilutions of the water samples; 0.5 mL of each dilution was inoculated into a tube containing 2.5 ml of alkaline

peptone water (APW, Oxoid, UK) supplemented with ampicillin (APW-A, 10 mg/L, Sigma-Aldrich, Steinheim, Germany). The tubes were incubated at 30 °C for 24 h. From each dilution, 100 μL was inoculated in ampicillin dextrin agar (ADA, CRITERION, Santa Monica, California, USA) plates and incubated at 30 °C for 24 h. When the ADA plates were found to be positive for the presence of *Aeromonas*, eight presumptive yellow colonies were subcultured in TSA and incubated under identical conditions. For molecular identifications performed at the genus level, the PCR method described by Chacón et al. (2002) was used. The MPN was calculated as described above.

***Legionella* spp. culture and typification** The detection and quantification of *Legionella* spp. in water samples were performed by culturing samples on BCYE agar supplemented with GVPC (MAIM, Spain). Samples (100 mL each) were filtered through a 0.45-μm pore size nylon membrane (Filter HNWP Millipore; Ireland), and the retained material was then suspended in 10 mL of Ringer 1/40 using vortexing for 2 min. The concentrates were cultured either directly or after two treatments: a thermal treatment at 50 °C for 30 min or an acid treatment in which 100 μL of acid buffer was added to 900 μL of the sample concentrate, as described in International Organization for Standardization (1998b).

Presumptive Legionella colonies were tested using a Legionella latex test (Oxoid, Basingstoke, Hampshire, England) according to the manufacturer's instructions. This kit enables the differentiation of *L. pneumophila* serogroup 1, *L. pneumophila* serogroups 2–14 and seven additional *Legionella* species.

***Legionella spp*. DNA extraction and qPCR analysis** Nucleic acids were extracted from 1 mL of Legionella sample concentrates using a Wizard genomic DNA purification kit (Promega, Madison, Wis.).

All samples were tested for the presence of *Legionella* spp. using a primer pair that was described in Herpers et al. (2003) and a probe described in Declerck et al. (2007). All reactions were performed in a final volume of 25 μL that contained 0.9 μM of each primer, 0.2 μM TaqMan MGB probe, 12.5 μL of 1× TaqMan Universal Master Mix and 5 μL of the extracted nucleic acids.

Legionella spp. positive samples were further tested for the presence of *L. pneumophila*. These assays targeted the *mip* gene and were based on the primers and probe described in Diederen et al. (2007). All of these reactions were performed in a final volume of 25 μL that contained 0.2 μM of the MipF primer, 0.3 μM of the MipR primer, 0.15 μM TaqMan probe Lpn-Mip, 12.5 μL of the 1× TaqMan Universal Master Mix and 5 μL of the extracted nucleic acids.

Free-living amoeba quantification

To quantify free-living amoebae (FLAs), 100 mL of each sample was concentrated to 10 mL by centrifugation at 800×*g* for 20 min. The concentrates were quantified by culturing them in non-nutrient agar (NNA) plates according to the MPN method described in Cervero-Aragó et al. (2013). MPN values were obtained from MPN tables (International Organization for Standardization 2005). The detection limit of the method ranged from 2×10^2 MPN of FLA/mL to 2×10^6 MPN of FLA/mL.

Statistical analyses

The data obtained for the concentrations of the microbiological parameters at the inlet and the outlet of the lagooning system were analysed using software packages developed in R (R Core Team 2013). The Wilcoxon test implemented in the R Package "exactRankTests" was used to calculate *p* values. This test was applied to calculate whether there were statistically significant differences between the values obtained at the inlet compared to those obtained at the outlet. The results relating to bacteria and free-living amoeba were analysed using two-sided Wilcoxon tests, while one-sided Wilcoxon tests were used to analyse results related to human viruses. The inclusion of the LOD for each microorganism and technique has been chosen to replace all non-detect. To compare the significance of the *p* values obtained for each microorganism, *p* values adjusted for multiple testing were calculated (Online resource 1) using the R package by applying the false discovery rate (FDR) test (Benjamini and Yekutieli 2001). A redefined adjusted *p* value of 0.05 was chosen to be the cut-off for statistical significance.

Pearson's correlation tests were performed using R software to determine if some relation existed between pH, water temperature, atmospheric temperature, theoretical hydraulic retention time and the microbial load at the outlet. At the same time, correlations among the different microorganisms at the outlet of the lagooning were tested.

Results

Results obtained for microbiological parameters in the inlet and outlet samples for each sampling point and date are shown in Table 2. Mean values before and after the lagooning tertiary treatment and the logarithmic reductions are presented in Tables 3 and 4 and Fig. 1. Negative or non-detected samples have been replaced by the LOD for each microorganism and technic. This conservative decision might underestimate the lagooning removal efficiency.

Table 2 Microbial quantification of virus, bacteria and protozoa analysed during the study period

	Sampling date	27/09/12	30/10/12	27/11/12	17/12/12	28/01/13	25/02/13	18/03/13	29/04/13	20/05/13	17/06/13	29/07/13	26/08/13
LI[a]	Human adenovirus (HAdV) (GC/100 mL)	1.48×10^2	4.50×10^3	9.01×10^4	1.61×10^4	6.91×10^3	ND[c]	8.27×10^4	2.45×10^4	2.49×10^4	1.73×10^3	1.05×10^3	3.78×10^3
LO[b]		8.54×10^1	ND[c]	ND[c]	1.33×10^3	1.69×10^3	ND[c]	1.59×10^4	3.72×10^4	ND[c]	6.16×10^2	ND[c]	ND[c]
LI[a]	JC polyomavirus (JCPyV) (GC/100 mL)	ND[c]	ND[c]	2.62×10^2	1.58×10^3	7.62×10^3	8.55×10^3	5.57×10^2	6.37×10^3	1.84×10^3	ND[c]	1.50×10^2	1.81×10^2
LO[b]		ND[c]	ND[c]	ND[c]	4.28×10^2	ND[c]	1.95×10^2	4.22×10^2	8.0×10^1	ND[c]	ND[c]	ND[c]	ND[c]
LI[a]	Human norovirus GI (NoVGI) (GC/100 mL)	2.75×10^4	ND[c]	ND[c]	1.19×10^5	1.42×10^5	2.41×10^4	5.32×10^4	7.56×10^3	2.70×10^4	2.98×10^3	ND[c]	ND[c]
LO[b]		1.28×10^3	ND[c]	ND[c]	ND[c]	4.04×10^5	2.09×10^4	2.50×10^4	1.19×10^4	4.50×10^2	2.72×10^3	ND[c]	ND[c]
LI[a]	Human norovirus GII (NoVGII) (GC/100 mL)	3.17×10^3	4.28×10^4	3.38×10^3	1.28×10^5	2.78×10^4	3.55×10^3	5.29×10^4	3.04×10^3	1.61×10^4	5.20×10^3	ND[c]	4.30×10^2
LO[b]		1.16×10^3	ND[c]	ND[c]	4.03×10^2	1.66×10^4	9.50×10^3	1.52×10^4	3.42×10^3	6.80×10^3	4.18×10^2	ND[c]	ND[c]
LI[a]	Heterotrophic bacteria 72 h at 22 °C (CFU/ 100 mL)	1.23×10^8	8.10×10^7	7.23×10^6	4.10×10^7	8.40×10^7	2.69×10^8	5.10×10^7	1.62×10^7	2.36×10^7	1.09×10^7	2.14×10^7	1.20×10^6
LO[b]		4.00×10^8	6.30×10^6	7.80×10^5	4.00×10^6	6.30×10^6	1.24×10^7	1.18×10^6	1.89×10^7	7.40×10^6	2.52×10^7	2.15×10^7	9.00×10^6
LI[a]	*Escherichia coli* (EC) (MPN/100 mL)	3.6×10^5	5.8×10^5	1.6×10^5	2.3×10^3	3.9×10^3	9.7×10^5	5.2×10^5	1.9×10^3	3.6×10^4	7.9×10^4	6.1×10^5	4.1×10^4
LO[b]		4.5×10^3	ND[c]	ND[c]	7.8×10^1	ND[c]	1.2×10^2	3.5×10^4	ND[c]	1.6×10^2	1.2×10^2	2.8×10^3	6.1×10^2
LI[a]	Intestinal enterococci (EI) (MPN/100 mL)	1.7×10^4	4.3×10^3	6.9×10^3	5.7×10^2	1.1×10^2	1.1×10^5	2.7×10^4	9.7×10^2	1.9×10^4	1.2×10^3	4.1×10^4	3.7×10^3
LO[b]		5.2×10^2	3.8×10^1	3.8×10^1	1.2×10^2	7.8×10^1	7.8×10^1	7.3×10^3	ND[c]	ND[c]	3.8×10^1	7.6×10^2	1.8×10^2
LI[a]	*Arcobacter* (MPN/100 mL)	1.1×10^8	2.3×10^6	2.4×10^7	2.4×10^7	1.07×10^7	2.3×10^7	1.27×10^7	1.23×10^5	$2.7E \times 10^7$	2.2×10^6	9.9×10^6	9.8×10^5
LO[b]		ND[c]	ND[c]	ND[c]	ND[c]	ND[c]	4.4×10^3	4.5×10^3	1.28×10^5	1.6×10^6	1.7×10^4	ND[c]	ND[c]
LI[a]	*Aeromonas* (MPN/100 mL)	+[d]	+[d]	+[d]	+[d]	+[d]	+[d]	1.4×10^6	2.6×10^5	3.3×10^6	9.1×10^5	4.3×10^5	1.6×10^5
LO[b]		+[d]	+[d]	+[d]	+[d]	+[d]	+[d]	4.3×10^3	6.5×10^3	4.2×10^4	3.4×10^5	3.4×10^5	4.3×10^5
LI[a]	*Legionella* spp. qPCR (GC/100 mL)	2.82×10^3	3.26×10^5	ND[c]	1.00×10^5	5.34×10^5	1.37×10^7	2.52×10^5	1.23×10^4	2.49×10^3	1.48×10^5	8.65×10^3	1.28×10^3
LO[b]		ND[c]	ND[c]	ND[c]	ND[c]	ND[c]	1.78×10^5	6.92×10^2	1.05×10^3	2.67×10^3	1.62×10^3	3.24×10^3	2.41×10^2
LI[a]	Free living amoeba (FLA) (MPN/100 mL)	1.1×10^6	4.9×10^5	2.2×10^4	1.7×10^5	2.3×10^4	7.9×10^4	4.9×10^4	4.9×10^4	1.7×10^6	1.4×10^6	1.7×10^4	2.2×10^4
LO[b]		1.4×10^4	2.3×10^4	4.9×10^4	4.9×10^5	8.0×10^3	9.0×10^3	2.0×10^3	3.3×10^5	3.3×10^5	2.3×10^4	1.3×10^4	1.7×10^4

[a] Lagooning inlet

[b] Lagooning outlet

Table 3 Inlet and outlet mean concentration of microbial parameters analysed and their logarithmic reduction after the lagooning treatment during the study period

Sampling period	Reference samples	Virus (CG/100 ml)							
		Human adenovirus (HAdV)		JC polyomavirus (JCPyV)		Norovirus GI (NoVGI)		Norovirus GII (NoVGII)	
		Geometric mean (min-max)	Positive samples	Geometric mean (min-max)	Positive samples	Geometric mean (min-max)	Positive samples	Geometric mean (min-max)	Positive samples
June–September	LI[a]	1.0×10^3 (1.05×10^3 –3.78×10^3)	4/4	1.15×10^2 (8.0×10^1 –1.81×10^2)	2/4	1.76×10^3 (3.43×10^2 –2.75×10^4)	2/4	1.13×10^3 (2.29×10^2 –5.20×10^3)	2/4
	LO[b]	8.19×10^1[c] (2.9×10^1 –6.16×10^2)	2/4	(8.0×10^1)	0/4	8.01×10^2 (3.43×10^2 –2.72×10^3)	2/4	3.99×10^2 (2.29×10^2 –1.16×10^3)	2/4
	Log reduction	1.09		2.06		0.34		0.45	
October–May	LI[a]	7.13×10^3 (2.9×10^1 –8.27×10^4)	7/8	1.39×10^3 (8.0×10^1 –8.55×10^2)	7/8	1.23×10^4 (3.43×10^2 –1.42×10^5)	6/8	1.62×10^4 (2.29×10^2 –1.28×10^5)	8/8
	LO[b]	4.21×10^2 (2.9×10^3 –3.72×10^4)	4/8	1.81×10^2 (8.0×10^1 –1.95×10^3)	3/8	3.82×10^3 (3.43×10^2 –4.04×10^5)	5/8	2.42×10^3 (2.29×10^2 –1.66×10^4)	7/8
	Log reduction	1.23		0.89		0.51		0.83	
Global	LI[a]	3.71×10^3 (1.48×10^2 –9.01×10^4)	11/12	6.06×10^2 (8.0×10^1 –7.62×10^3)	9/12	6.43×10^3 (3.43×10^2 –1.42×10^5)	8/12	6.67×10^3 (2.29×10^2 –1.28×10^5)	11/12
	LO[b]	2.44×10^2 (2.9×10^1 –3.72×10^4)	6/12	1.38×10^2 (8.0×10^1 –1.95×10^3)	3/12	2.27×10^3 (3.43×10^2 –4.04×10^5)	7/12	1.28×10^3 (2.29×10^2 –1.66×10^4)	8/12
	Log reduction	1.18[d]		0.64[d]		0.45		0.72	

Geometric means are calculated by using the LOD for a given microorganism and technique. The minimum and maximum value for each microorganism is specified

[a] Lagooning inlet

[b] Lagooning outlet

[c] No positive values observed

[d] Statistically significative p value 0.05

In general, all studied microorganisms were reduced in concentration after lagooning, with reductions ranging from 1.18 to 0.45 log for some enteric viruses and up to more than 2 log for EC (Fig. 1). The observed reductions in human viral faecal markers (HAdV and JCPyV) and FIB (EC and IE) were statistically significant (Online

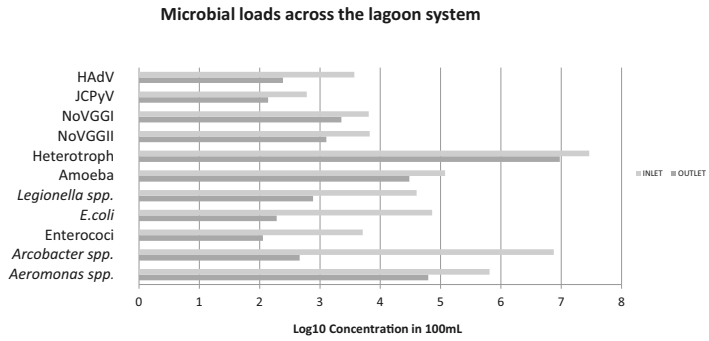


Fig. 1 Mean microbial concentration loads of the inlet and outlet reclaimed water produced in the lagooning system studied

resource 1). No significant statistical correlation was found between the measured physico-chemical water parameters and any of the microorganisms analysed at the lagooning outlet. A significative negative correlation was found between HRT and heterotrophic bacterial counts ($R^2 = -0.72$, p value = 0.0078). Only significative correlations between norovirus genogroups GI and GII ($R^2 = 0.84$, p value = 0.0006) were found.

The average water retention time in the system was estimated to be of 31 days with two main periods: a high demand period (from June to September 25,000 m^3/month) with an average estimated retention time of 16 days and a low demand period (from October to May, 25,000 m^3/month) with an average estimated retention time of 88 days (see Tables 3 and 4). Retention time was influenced by raining events, which decreased the water demand from farmers. Although initial analyses appear to indicate differences in microbial removal between these periods, the small number of samples tested for each period group does not enable to conclude if differences are statistically significant.

Efficiency of virus removal and inactivation

A high abundance of viruses was observed in the secondary effluents analysed over the entire year. The majority of the samples (11/12) entering into the lagooning system were positive for HAdV with a mean value of 3.71×10^3 GC/100 mL, whereas only half of the outlet samples were positive (6/12), with mean values of 2.44×10^2 GC/100 mL (Tables 2 and 3). For JCPyV, the number of positive samples was high (9/12) at the inlet, with a mean value of 6.06×10^2 GC/100 mL, while the number of positive samples decreased at the outlet (3/12), where samples showed a mean value of 1.38×10^2 GC/100 mL. Human pathogens NoV GI and GII were highly prevalent at the inlet (in 8/12 and 11/12 of the samples, respectively). The concentrations of both genogroups at the inlet of the system were 6.43×10^3 GC/100 mL and 6.67×10^3 GC/100 mL, respectively, while the prevalence of these viruses at the outlet of the system slightly decreased (7/12 of the samples at 2.27×10^3 GC/100 mL and 8/12 of the samples at 1.38×10^3 GC/100 mL, respectively). None of the inlet or outlet samples were positive for HEV.

The infectivity experiments presented one out of four samples with infectious HAdV in the inlet (233 MPN IU/100 mL). None of the four outlet samples tested showed infectious HAdV (LOD 8MPN infective virus/100 mL). The positive inlet sample for HAdV was amplified, and the amplicon obtained was sequenced and corresponded to HAdV41 (nucleotide accession number HG976918).

Efficiency of inactivation of HBC and standard FIB

HBC

Heterotrophic bacteria were present in all of the samples analysed, with a geometric mean of 2.91×10^7 cfu/100 mL in the inlet samples and 9.43×10^6 cfu/100 mL in the outlet samples. The results showed a reduction of 0.49 logs in the HBC count, but this difference was not statistically significant (Table 4 and Online resource 1). Moreover, these differences varied slightly during the year. A peak reduction in HBC of 0.95 logs was observed during the months with lower water demand. In contrast, there was an increase in the bacterial count to 0.44 logs in the summer during the period of high water demand.

Standard FIB

All 12 inlet water samples were positive for EC and IE, with geometric means of 7.23×10^4 MPN/100 mL and 5.11×10^3 MPN/100 mL, respectively (Table 2). Only eight outlet water samples were positive for EC (1.92×10^2 MPN/100 mL), whereas 10 outlet samples were positive for IE (1.14×10^2 MPN/100 mL). The lagooning reduced 2.58 and 1.65 log EC and IE, respectively (Table 4).

During the low demand period, the mean concentration of EC in the inlet water was 4.81×10^4 MPN/100 mL, while in the outlet, the mean concentration was 8.47×10^1 MPN/100 mL. These data indicate a 2.75 log reduction. For the same period, the mean concentration of IE was 4.23×10^3 MPN/100 mL for the inlet water and 8.03×10^1 MPN/100 mL for the outlet water, with an observed reduction of 1.72 logs.

During the high demand period, the mean concentration of EC was 1.63×10^5 MPN/100 mL and 9.80×10^2 MPN/100 mL for the inlet and the outlet water, respectively, indicating a 2.22 log reduction. During this period, the concentration of IE in the inlet water was 7.46×10^3 MPN/100 mL, while the concentration at the outlet was 2.28×10^2 MPN/100 mL, indicating a 1.51 log reduction.

Evaluation of the potential regrowth of bacteria and other opportunistic pathogens

Arcobacter

All 12 inlet water samples were positive for *Arcobacter*, while only 5 (41.6 %) of the outlet water samples were positive (Table 2). The average concentration of *Arcobacter* in the inlet water samples was 7.51×10^6 MPN/100 mL, while in the outlet water samples, the average concentration was 4.59×10^2 MPN/100 mL (Table 4). During the low demand period, the mean concentration of *Arcobacter* in the inlet water was

Table 4 Inlet and outlet mean concentrations and their logarithmic reduction after the lagooning treatment during the study period

Sampling period	Water source	Heterotrophs (CFU/100 ml)		FIB (MPN/100 ml)				*Arcobacter* spp. (MPN/100 ml)
				EC		EI		
		Geometric mean (min-max)	Positive samples	Geometric mean (min-max)	Positive samples	Geometric mean (min-max)	Positive samples	Geometric mean (min-max)
June–September (high demand)	LI[a]	1.36×10^7 (1.20×10^6–1.23×10^8)	12/12	1.63×10^5 (4.10×10^4–6.10×10^5)	4/4	7.46×10^3 (1.20×10^3–4.10×10^4)	4/4	6.96×10^6 (9.80×10^5–1.10×10^8)
	LO[b]	3.74×10^7 (9.0×10^6–4.0×10^8)	12/12	9.80×10^2 (1.20×10^2–4.50×10^3)	4/4	2.28×10^2 (3.80×10^1–7.60×10^2)	4/4	(1.70×10^4)
	Log reduction	−0.44 (regrowth)		2.22		1.51		4.81
October/May (low demand)	LI[a]	4.25×10^7 (7.23×10^6–2.69×10^8)	12/12	4.81×10^4 (2.90×10^3–9.70×10^5)	8/8	4.23×10^3 (1.10×10^2–1.10×10^5)	8/8	7.80×10^6 (1.23×10^5–2.70×10^7)
	LO[b]	4.74×10^6 (7.80×10^5–1.89×10^7)	12/12	8.47×10^1 (1.5×10^1–3.50×10^4)	4/8	8.03×10^1 (1.5×10^1–7.30×10^3)	6/8	9.47×10^2 (2.0×10^1–1.60×10^6)
	Log reduction	0.95		2.75		1.72		3.92
Global	LI[a]	2.91×10^7 (1.20×10^6–4.0×10^8)	12/12	7.23×10^4 (2.90×10^3–9.70×10^5)	12/12	5.11×10^3 (1.1×10^2–1.1×10^5)	12/12	7.51×10^6 (9.80×10^5–1.10×10^8)
	LO[b]	9.43×10^6 (7.80×10^5–4.0×10^8)	12/12	1.92×10^2 (1.5×10^1–3.50×10^4)	8/12	1.14×10^2 (1.5×10^1–7.30×10^3)	10/12	4.59×10^2 ($2.0 \times 10^1 \times 10^3$–$1.60 \times 10^6$)
	Log reduction	0.49		2.58[c]		1.65[c]		4.21[c]

Sampling period	*Arcobacter* spp. (MPN/100 ml)	*Aeromonas* spp. (MPN/100 ml)		*Legionella* spp. (GC/100 mL)		Free-living amoeba (MPN/100 ml)	
	Positive samples	Geometric mean (min-max)	Positive samples	*Legionella* spp. (GC/100 mL)	Positive samples	Geometric mean (min-max)	Positive samples
June–September (high demand)	4/4	3.97×10^5 (1.60×10^5–9.10×10^5)	3/3	7.22×10^4 (1.28×10^3–1.48×10^5)	4/4	7.21×10^5 (1.70×10^4–2.20×10^4)	12/12
	4/4	3.68×10^5 (3.40×10^5–4.30×10^5)	3/3	7.63×10^3 (1.0×10^2–1.62×10^4)	3/4	4.50×10^3 (1.30×10^4–2.30×10^4)	12/12
	4.81	0.03		0.89		0.98	
October/May (low demand)	8/8	1.06×10^6 (2.60×10^5–3.30×10^6)	3/3	6.07×10^4 (1.0×10^2–1.37×10^7)	7/8	1.04×10^5 (2.20×10^4–1.70×10^6)	12/12
	4/8	1.05×10^4 (4.30×10^3–4.20×10^4)	3/3	6.56×10^2 (6.92×10^2–1.78×10^5)	4/8	4.14×10^4 (6.92×10^2–1.78×10^5)	12/12
	3.92	2.0		1.97		0.40	
Global	12/12	1.08×10^6 (1.60×10^5–3.30×10^6)	6/6	3.12×10^4 (1.0×10^2–1.37×10^7)	11/12	1.19×10^5 (2.20×10^4–1.70×10^6)	12/12
	5/12	1.94×10^5 (4.30×10^3–4.30×10^5)	6/6	7.70×10^2 (2.41×10^2–1.78×10^5)	6/12	3.04×10^4 (6.92×10^2–1.78×10^5)	12/12
	4.21[c]	1.02		1.72		0.59	

Geometric means are calculated by using the LOD for a given microorganism and technique. The minimum and maximum value for each microorganism is specified

[a] Lagooning inlet

[b] Lagooning outlet

[c] Statistically significative p value 0.05

7.80×10^6 MPN/100 mL, while in the outlet, it was 9.47×10^2 MPN/100 mL. This data represents a 3.92 log reduction. Notice that during the high demand period, the mean concentration of *Arcobacter* was 6.9×10^6 MPN/100 mL for inlet water and 1.08×10^2 MPN/100 mL for outlet water, and these data represent a 4.81 log reduction.

Aeromonas spp.

All water samples (12 inlet and 12 outlet) were positive for *Aeromonas* (Table 2). Over the 6 months during which quantification was performed (March 2013 and August 2013), the average concentration of *Aeromonas* spp. in the inlet water was 1.08×10^6 MPN/100 mL and the average concentration in the outlet was 1.94×10^5 MPN/100 mL. These data represent a 1.02 log reduction in the *Aeromonas* spp. load during the storage period (Table 4). During the lower demand period, the mean concentration of *Aeromonas* in the inlet water was 1.06×10^6 MPN/100 mL, while in the outlet water, it was 1.05×10^4 MPN/100 mL. These data indicate a 2 log reduction. In contrast, during the high demand period, the mean concentration of *Aeromonas* was 3.97×10^5 MPN/100 mL in the inlet water and 3.68×10^5 MPN/100 mL in the outlet water, representing only 0.03 log reduction.

Legionella spp.

Legionella spp. were detected using qPCR, but not culture methods, in both inlet and outlet samples. In the samples with higher concentrations, which were obtained from February to April, it was not possible to isolate any *Legionella* spp. because there was a high concentration of accompanying microbiota that also grew on the BCYE agar plates. Hence, it was not possible to differentiate between species or serogroups using sero-agglutination. The qPCR results showed that *Legionella* spp. were present in 11 out of the 12 samples obtained from the lagooning inlet and in 7 out of 12 samples obtained from the outlet (Table 2). Within these samples, *L. pneumophila* was detected in six inlet samples and two outlet samples. Overall, a global reduction of 1.72 log was observed in *Legionella* (Table 4). However, when only the positive samples obtained during the low and high water demand periods were compared, we observed a log reduction of 1.97 and 0.89, respectively, in *Legionella*.

FLA

All inlet and outlet water samples contained FLA, with geometric means of 1.19×10^5 MPN/100 mL and 3.04×10^4 MPN/100 mL, respectively. In general, higher concentrations of FLA were detected in the inlet samples, and a reduction of 0.59 log was observed after lagooning. Moreover, an 0.40 log reduction was observed during the low water demand period, and a 0.98 log reduction was observed during the high demand.

Discussion

In this work, a lagooning system, considered as a natural and low-cost tertiary disinfection method, is used to treat the secondary wastewater effluent produced by a wastewater treatment plant.

The stabilization pond system studied in this report decreased the concentration of all microorganisms analysed, with the exception of FLA and HBC which presented on some months similar counts at the outlet. Higher numbers of HBC in the outlet samples were observed during periods with higher water temperatures, despite increased solar radiation. These findings support the idea that temperature can positively impact the regrowth of HBC, as was reported by Niquette et al. (2001). Microorganisms that are able to grow in aquatic environments, such as some of the studied bacteria and protozoa, were analysed for either regrowth or inactivation. It is known that lagooning applied to produce reclaimed water reduces the levels of bacteria in the effluent (Jjemba et al. 2010; Derry and Attwater 2014). Nevertheless, bacteria regrowth has been observed in the mentioned studies in the reservoir and distribution systems, where there was a loss of residual disinfectant and high levels of assimilable organic carbon. Despite this fact, HBC and FLA should be considered when reclaimed water produced by lagooning is going to be used for irrigation purposes as both groups contain potentially pathogenic microorganisms and their regrowth may represent a health risk.

Viral concentrations decreased between 0.45 and 1.18 log (NoVGI and HAdV respectively) (Table 3). The obtained values were similar to those reported by Maynard et al. (1999); however, depending on the lagooning characteristics, higher reductions can be reached (up to 2 log) (Shuval 1990; Pay Drechsel et al. 2010), depending on the characteristics of the lagooning system. Jurzik et al. (2015) reported a high degree of effectiveness in the removal of FIB and bacteriophages in a lagooning system, whereas human viruses, such as HAdV and JCPyV, were not significantly decreased. The lack of infectivity assays in previous studies constitutes an important limitation as those viral particles detected could not be infectious. In the current study, higher reductions of genome copies were observed for the same faecal virus markers. This might be related to the higher water temperatures expected from a Mediterranean site and also due to differences in the retention times. Infectivity results showed that one out of four inlet samples but none of the four outlet samples contained infective HAdV particles (less than 8 MPN infectious HAdV in the 100 mL of sample analysed), showing that lagooning could achieve a reduction of 1.49 log. It is widely accepted

that cell culture assays have some limitations, including the bias produced by the efficiency or inability of certain adenovirus genotypes to replicate in specific cell lines. Hence, these data should be taken only as an indication of the risk of infection.

Noroviruses are the leading cause of foodborne disease outbreaks worldwide (Koo et al. 2010), most of those sporadic cases and outbreaks being related to NoVGII (Lopman et al. 2004; Lodder and de Roda Husman 2005; Kroneman et al. 2008). This virus was prevalent in the inlet samples throughout the year and showed a seasonal peak in winter, from December to May, as previously reported (Haramoto et al. 2006; Katayama et al. 2008; Nordgren et al. 2009), with higher viral titres ranging from 10^4 to 10^5 GC/100 mL. July was the only month during which NoVGII was not detected. NoVGI was the less prevalent of the two genogroups, but when it was present, its viral titres were higher than those of NoVGII (Table 2). Higher resistance to wastewater treatment was observed for NoVGI compared to NoVGII, as previously reported (Da Silva et al. 2007; Nordgren et al. 2009). Norovirus titres should be a matter of concern, especially considering its low infectious dose of 18 particles (Teunis et al. 2008). The risk of infection through the consumption of raw edible vegetables irrigated with reclaimed water, containing NoV genomic copies, has been recently quantified (Sales-Ortells et al. 2015).

HEV was not found in either the inlet or the outlet samples of the lagooning system, even though it is widely known that this virus circulates in industrialized countries (Legrand-Abravanel et al. 2009; Masclaux et al. 2013). The low millilitre equivalents (8.75 mL) of the samples tested by molecular methods when samples are concentrated by skimmed milk flocculation (Rusiñol et al. 2015) in combination with the lower prevalence of HEV virus compared to other enteric viruses (Masclaux et al. 2013) may explain the absence of positive results.

The viral reduction values reported by Jurzik et al. (2015), in addition to those in the current study, indicate that the use of traditional FIB and bacteriophages as surrogates for predicting the presence or absence of viral pathogens in reclaimed water is not always reliable (Baggi et al. 2001; Hot et al. 2003; Ottoson et al. 2006a). The lack of correlation found between FIB and viral faecal markers at the lagooning outlet reinforces that idea, as previously reported (Mara and Pearson 1987; Donnison and Ross 1995).

A reduction of more than two logs was observed for EC in the lagooning system. This value is slightly superior to the average removal value reported for EC by Goyal (2013). In the case of IE, although similar removal load was observed, this indicator appears to be more resistant to outdoor storage than EC because only two samples were below the detection limit for IE, whereas four of the outlet samples were below the limit for EC. The higher survival capacity of IE compared to EC is well described (Fleischer et al. 2000; Figueras and Borrego 2010), and the results obtained in this study are in agreement with those reported (Tyagi et al. 2008). In Spain, reclaimed water is controlled under the regulation RD 1620/2007, which establishes water uses according to different EC levels. The removal efficiency achieved in the lagooning system in this study was not enough to meet the regulation requirements because on some occasions the concentration of EC was higher than 100 MPN/100 mL which is the maximum allowed by the legislation for raw edible vegetables (RD 1620/2007). Therefore, the produced reclaimed water would not be suitable for some irrigation purposes (e.g., to irrigate raw-edible vegetables). Specifically, the outlet water exceeded the 100 CFU/100 mL recommended for EC in 7 out of 12 of the samples tested. Moreover, three out of seven of the positive outlet water samples had concentration values that were higher than the 1000 CFU/100 mL faecal coliforms. These higher values were observed also in summer during the time of the year when lagooning water is mainly used for irrigation.

Arcobacter which was present at high concentrations (7.51×10^6 MPN/100 mL) at the inlet has been clearly reduced 4 log (p value 6.21×10^{-4}). In fact, four out of the seven samples in which *Arcobacter* was not detected were also negative for the viruses tested. This indicates that lagooning is very effective in removing some potentially pathogenic bacteria. Numerous studies have shown that *Arcobacter* is abundant in wastewater. No other studies have systematically quantified the presence of *Arcobacter* in wastewater or its presence in a lagooning system as we have done in this study. In previous studies performed by our group, a correlation was demonstrated between the presence of *Arcobacter* and the presence of faecal pollution (Collado et al. 2008, 2010; Fisher et al. 2014). However, recently, in a study that used metagenomics, it was shown that the abundance of *Arcobacter* is due to its growth within the sewer environment and not due to human input. This conclusion was based on the low abundance of these microbes in the faeces of symptomatic and asymptomatic patients with diarrhoea (Figueras et al. 2014; Fisher et al. 2014). Other authors have reported the presence of *Arcobacter* spp. in sewage in the UK and associated its detection with the underestimation of these bacteria in the human community (Merga et al. 2014). Despite this, underestimation continues to occur, as we have suggested in previous studies (Collado and Figueras 2011; Fisher et al. 2014). These facts alone do not explain the high concentration found in sewage, from which *Arcobacter* can be isolated by direct plating without any enrichment. This latter finding indicates the growth and amplification capacity of these bacteria in sewage.

The low reduction (0.03 log) observed in *Aeromonas* during the warm season (June–September) correlated with high temperatures that ranged between 23 and 28 °C

(corresponding to the optimum growth temperature for this bacteria) and also with the low retention time of the water in the lagooning system (Table 4). These results agree with those of Monfort and Baleux (1990), who studied the *Aeromonas* dynamics in a sewage treatment pond and reported a slightly higher reduction in winter (99.8 %) than in summer (98.3 %). These authors found a positive correlation between pond water temperature and *Aeromonas* concentrations. However, opposite results were reported by (Hassani et al. 1992) in a study performed in Morocco, where the removal efficacy of the stabilization pond treatment used to clear domestic wastewater was higher in the warm months (98.8 %), when temperatures were approximately 30 °C, than in the colder months (97 %), when the temperatures were lower than 21 °C.

Isolating *Legionella* from complex environmental samples is a well-known arduous job (Joly et al. 2006; Serrano-Suárez and Araujo 2013; Blanky et al. 2015). The difficulties involved in isolating *Legionella* using culture methods include the low sensitivity of the culture media wherein other bacteria with faster growth rates than *Legionella* spp. suppress or mask its growth, or the fact that under stressful conditions, *Legionella* spp. enter into a viable but non-culturable (VBNC) stage (Steinert et al. 1997; Joly et al. 2006; Rodríguez-Martínez et al. 2015). However, the use of molecular techniques, such as qPCR, has enabled the detection of similar *Legionella* concentrations, as have been described in the literature (Palmer et al. 1993; Medema et al. 2004; Declerck et al. 2007). *Legionella* was more abundant in the inlet samples than in the outlet samples (Table 2). Nevertheless, a more substantial reduction was observed in winter, when the retention time was longer and temperatures were lower (Table 4). In summer, the lower reduction might be explained by the shorter retention time and higher temperatures, which were closer to the replication temperature of the bacteria. Half of the inlet samples contain *L. pneumophila*, but it was detected only in two of the outlet samples (data not shown). Hence, according to Spanish legislation (RD 1620/2007), this water can be considered as safe water and used for aerosoled irrigation. However, further studies are needed to improve the methods recommended for the detection of this bacteria in complex water matrices and also its health risk associated to their presence in reclaimed irrigation water.

Free-living amoeba were detected in all of the samples analysed, with no clear trend related to water matrix, retention time, temperature or other physicochemical parameters. The FLAs are a complex and heterogeneous group of microorganisms characterized by having two live stages: trophozoites and cysts. The cysts, which are a resistant and dormant form, enable these species to survive harsh environmental conditions, including the disinfection methods used in conventional WWTP. The presence of FLA in the effluents of WWTPs has been reported in several studies (Garcia et al. 2013; Magnet et al. 2013). The high numbers of trophozoites observed in the inlet and outlet water samples suggest that FLAs play a key role in the lagooning ecosystem. These species increase the presence of some microorganisms rather than others by acting as grazers (Danes and Cerva 1981, 1984; Greub and Raoult 2004; Lorenzo-Morales et al. 2007; Declerck 2010). This fact has promoted adaptative changes in the microorganisms in which they prey on, which must survive amoebal predation. For example, the presence of FLA may explain the presence of *Legionella* spp. because these bacteria have the ability to replicate within FLA that play a protective role against harsh environmental conditions (Richards et al. 2013; Cervero-Aragó et al. 2015). Recently, a publication reported the internalization of HAdV by ciliates in wastewater (Battistini et al. 2013). As has been previously shown in some bacterial species, viruses may use FLA or other protozoa as a shelter to wastewater treatments (Scheid and Schwarzenberger 2012). In the current study, no significant correlation was found between HAdV and FLA in the outlet samples. Further studies of these species could improve our understanding of viral survival in the environment. Unfortunately, the methodology used in the current study did not enable the identification of the genera of the isolated FLA, but data obtained provides an overview of its presence in two different water matrixes.

No significant regrowth of opportunistic pathogens was observed throughout the lagooning system. The microorganism removal efficiency of other tertiary waste water disinfection methods, such as membrane filtration, has been shown to reduce human noroviruses by less than 1 log and to reduce EC and IE by of 3.23 and 3.17 logs, respectively (Ottoson et al. 2006b). Chlorination after secondary treatment reduced IE by up to 2.57 logs and EC 1.18 logs, whereas HAdV was reduced by 0.81 logs (Francy et al. 2012). UV irradiation (254 nm) reduced EC, IE and human adenoviruses by 3.82, 3.32 and 0.24 logs, respectively, at a genomic copy level (Francy et al. 2012). The application of a more complex wastewater treatment method that is composed of membrane ultrafiltration in combination with chlorination and UV disinfection reduced HAdV (qPCR data), EC and IE by 1.44, 2.12 and 1.84 logs, respectively (Rusiñol et al. 2015). Similar logarithmic removal values of HAdV–at a genomic copy level—EC and IE were obtained in the lagooning in comparison with conventionally tertiary water treatments. However, the high construction costs and maintenance of more complex systems in combination with the difficulty of applying them in low-income developing countries make lagooning a sustainable and effective method of producing reclaimed water for irrigation purposes. Nevertheless, a quality control system should be implemented to ensure that the reclaimed water requirements contained in the legislation are fulfilled.

Conclusions

1. The results obtained in the present study demonstrate the variability of removal efficiency in lagooning systems as previously reported by other authors (Berg 1973; Maynard et al. 1999). The lagooning system evaluated in this study achieved significant logarithmic reductions in the human viral faecal markers HAdV and JCPyV ranging from 1.18 (0.24–2.93) to 0.64 (0.12–1.97), at the genomic copy level respectively. A 2.58 (1.17–4.59) and 1.65 (0.15–3.14) EC and IE log reduction was observed.
2. No regrowth of FIBs was observed in the system, which obtained a reduction of nearly 2 logs between the inlet and outlet samples. However, although FIBs were reduced, in half of the samples analysed, the concentration at the outlet exceeded the recommendations of the Spanish legislation (RD/1620/2007). The absence of FIBs does not guarantee the absence of viruses because some samples that were negative for FIBs presented viral faecal markers. Therefore, the inclusion of viral faecal markers, such as HAdV and JCPyV, in reclaimed water legislation should be considered to minimize risks.
3. Opportunistic pathogens, common inhabitants of water systems as *Legionella* spp. and *Aeromonas* spp., showed a pattern of reduction that was different from that of FIBs, while the pattern observed for *Arcobacter* was more in agreement with that of FIBs. In addition, HBC and FLA, which are microorganisms that are representative of complex heterogeneous groups, showed small reductions throughout the lagooning, and in some occasions, their counts were higher in samples collected at the outlet than in those collected at the inlet, suggesting bacterial and protozoa regrowth. This fact reinforces the existence of two different microbial communities. Both communities are differently influenced by environmental factors such as temperatures above 20 °C, but further studies are necessary to confirm these trends and to obtain a better understanding of the composition of these populations.
4. Data obtained from this study reinforces the idea that more studies on lagooning systems are required for improving its design and management in order to fulfil the safety requirements established in the RD 1620/2007 and ensure the production of safe reclaimed water to irrigate raw edible vegetables.

Acknowledgments The study reported here was partially founded by programa Recercaixa 2012 ACUP-00300 and AGL2011-30461-C02-01/ALI of the Spanish Ministry of Science and Innovation and Consolidated Research Groups Program of the Catalan Government (2014SGR529 and VirBaP 2014SRG914). Xavier Fernandez-Cassi and Fadua Latif-Eugenín acknowledge the AGAUR Agency of the Catalan Government and the Universitat Rovira i Virgili, respectively, which supported their PhD grants. The authors would like to thank the lagooning system manager, Conrad Martorell, for his collaboration during this project.

References

Allard A, Albinsson B, Wadell G (2001) Rapid typing of human adenoviruses by a general PCR combined with restriction endonuclease analysis. J Clin Microbiol 39(2):498–505, Available from: http://jcm.asm.org/content/39/2/498

Baggi F, Demarta A, Peduzzi R (2001) Persistence of viral pathogens and bacteriophages during sewage treatment: lack of correlation with indicator bacteria. Res Microbiol 152(8):743–51, Available from: http://www.ncbi.nlm.nih.gov/pubmed/11686388

Battistini R, Marcucci E, Verani M, Di Giuseppe G, Dini F, Carducci A (2013) Ciliate-adenovirus interactions in experimental co-cultures of Euplotes octocarinatus and in wastewater environment. Eur J Protistol 49(3):381–8, Available from: http://www.sciencedirect.com/science/article/pii/S0932473912000946

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat 29(4):1165–88

Berg G (1973) Removal of viruses from sewage, effluents, and waters. I. A review. Bull World Health Organ 49(5):451–60, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2480995&tool=pmcentrez&rendertype=abstract

Bernard H, Faber M, Wilking H, Haller S, Höhle M, Schielke A et al (2014) Large multistate outbreak of norovirus gastroenteritis associated with frozen strawberries, Germany, 2012. Euro Surveil 19(8):20719, Available from: http://www.ncbi.nlm.nih.gov/pubmed/24602278

Blanky M, Rodríguez-Martínez S, Halpern M, Friedler E (2015) Legionella pneumophila: from potable water to treated greywater; quantification and removal during treatment. Sci Total Environ 533:557–65, Available from: http://www.ncbi.nlm.nih.gov/pubmed/26188406

Bofill-Mas S, Albinana-Gimenez N, Clemente-Casares P, Hundesa A, Rodriguez-Manzano J, Allard A et al (2006) Quantification and stability of human adenoviruses and polyomavirus JCPyV in wastewater matrices. Appl Environ Microbiol 72(12):7894–6, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1694247&tool=pmcentrez&rendertype=abstract

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 Res Int 2013:192089, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3671278&tool=pmcentrez&rendertype=abstract

Calgua B, Mengewein A, Grunert A, Bofill-Mas S, Clemente-Casares P, Hundesa A et al (2008) Development and application of a one-step low cost procedure to concentrate viruses from seawater samples. J Virol Methods 153(2):79–83, Available from: http://www.ncbi.nlm.nih.gov/pubmed/18765255

Calgua B, Rodriguez-Manzano J, Hundesa A, Suñen E, Calvo M, Bofill-Mas S et al (2013) New methods for the concentration of viruses from urban sewage using quantitative PCR. J Virol Methods 187(2):215–21, Available from: http://www.sciencedirect.com/science/article/pii/S0166093412003710

Callejón RM, Rodríguez-Naranjo MI, Ubeda C, Hornedo-Ortega R, Garcia-Parrilla MC, Troncoso AM (2015) Reported foodborne outbreaks due to fresh produce in the United States and European Union: trends and causes. Foodborne Pathog Dis 12(1):32–8, Available from: http://www.ncbi.nlm.nih.gov/pubmed/25587926

Campos C, Guerrero A, Cárdenas M (2002) Removal of bacterial and viral faecal indicator organisms in a waste stabilization pond system in Choconta, Cundinamarca (Colombia). Water Sci Technol 45(1):61–6, Available from: http://www.ncbi.nlm.nih.gov/pubmed/11833732

Carter MJ (2005) Enterically infecting viruses: pathogenicity, transmission and significance for food and waterborne infection. J Appl

Microbiol 98(6):1354–80, Available from: http://www.ncbi.nlm.nih.gov/pubmed/15916649

Cervero-Aragó S, Rodríguez-Martínez S, Canals O, Salvadó H, Araujo RM (2013) Effect of thermal treatment on free-living amoeba inactivation. J Appl Microbiol. Oct 30 [cited 2014 Dec 28]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/24251398

Cervero-Aragó S, Rodríguez-Martínez S, Puertas-Bennasar A, Araujo RM (2015) Effect of common drinking water disinfectants, chlorine and heat, on free legionella and Amoebae-associated Legionella. PLoS One 10(8):e0134726, Available from: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0134726

Chacón MR, Castro-Escarpulli G, Soler L, Guarro J, Figueras MJ (2002) A DNA probe specific for Aeromonas colonies. Diagn Microbiol Infect Dis 44(3):221–5, Available from: http://www.ncbi.nlm.nih.gov/pubmed/12493167

Collado L, Figueras MJ (2011) Taxonomy, epidemiology, and clinical relevance of the genus Arcobacter. Clin Microbiol Rev 24(1):174–92, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3021208&tool=pmcentrez&rendertype=abstract

Collado L, Inza I, Guarro J, Figueras MJ (2008) Presence of Arcobacter spp. in environmental waters correlates with high levels of fecal pollution. Environ Microbiol 10(6):1635–40, Available from: http://www.ncbi.nlm.nih.gov/pubmed/18215159

Collado L, Kasimir G, Perez U, Bosch A, Pinto R, Saucedo G et al (2010) Occurrence and diversity of Arcobacter spp. along the Llobregat River catchment, at sewage effluents and in a drinking water treatment plant. Water Res 44(12):3696–702, Available from: http://www.ncbi.nlm.nih.gov/pubmed/20427071

R Core Team (2013) R: A language and environment for statistical computing. [Internet]. Available from: http://www.r-project.org/

Da Silva AK, Le Saux JC, Parnaudeau S, Pommepuy M, Elimelech M, Le Guyader FS (2007) Evaluation of removal of noroviruses during wastewater treatment, using real-time reverse transcription-PCR: different behaviors of genogroups I and II. Appl Environ Microbiol 73(24):7891–7

Danes L, Cerva L (1981) Survival of polioviruses and echoviruses in Acanthamoeba castellanii cultivated in vitro. J Hyg Epidemiol Microbiol Immunol 25(2):169–74, Available from: http://www.ncbi.nlm.nih.gov/pubmed/6265550

Danes L, Cerva L (1984) Poliovirus and echovirus survival in Tetrahymena pyriformis culture in vivo. J Hyg Epidemiol Microbiol Immunol 28(2):193–200, Available from: http://www.ncbi.nlm.nih.gov/pubmed/6088623

Dare A (2015) Irrigation with treated municipal wastewater in Indiana, United States. J Soil Water Conserv 70(4):89A–94A, Available from: http://www.jswconline.org/content/70/4/89A.extract

Declerck P (2010) Biofilms: the environmental playground of Legionella pneumophila. Environ Microbiol 12(3):557–66, Available from: http://www.ncbi.nlm.nih.gov/pubmed/19678829

Declerck P, Behets J, van Hoef V, Ollevier F (2007) Detection of Legionella spp. and some of their amoeba hosts in floating biofilms from anthropogenic and natural aquatic environments. Water Res 41(14):3159–67, Available from: http://www.ncbi.nlm.nih.gov/pubmed/17544473

Derry C, Attwater R (2014) Regrowth of enterococci indicator in an open recycled-water impoundment. Sci Total Environ 468-469:63–7, Available from: http://www.ncbi.nlm.nih.gov/pubmed/24008073

Diederen BMW, de Jong CMA, Marmouk F, Kluytmans JAJW, Peeters MF, Van der Zee A (2007) Evaluation of real-time PCR for the early detection of Legionella pneumophila DNA in serum samples. J Med Microbiol 56(Pt 1):94–101, Available from: http://www.ncbi.nlm.nih.gov/pubmed/17172523

Donnison AM, Ross CM (1995) Somatic and F-specific coliphages in New Zealand waste treatment lagoons. Water Res 29(4):1105–10, Available from: http://www.sciencedirect.com/science/article/pii/004313549400227X

Erker JC, Desai SM, Schlauder GG, Dawson GJ, Mushahwar IK (1999) A hepatitis E virus variant from the United States: molecular characterization and transmission in cynomolgus macaques. J Gen Virol 80(3):681–90, Available from: http://www.ncbi.nlm.nih.gov/pubmed/10092008

Figueras MJ, Borrego JJ (2010) New perspectives in monitoring drinking water microbial quality. Int J Environ Res Public Health 7(12):4179–202, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3037048&tool=pmcentrez&rendertype=abstract

Figueras MJ, Levican A, Collado L (2012) Updated 16S rRNA-RFLP method for the identification of all currently characterised Arcobacter spp. BMC Microbiol 12:292, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3548738&tool=pmcentrez&rendertype=abstract

Figueras MJ, Levican A, Pujol I, Ballester F, Rabada Quilez MJ, Gomez-Bertomeu F (2014) A severe case of persistent diarrhoea associated with Arcobacter cryaerophilus but attributed to Campylobacter sp. and a review of the clinical incidence of Arcobacter spp. New Microbes New Infect 2(2):31–7, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4184587&tool=pmcentrez&rendertype=abstract

Fisher JC, Levican A, Figueras MJ, McLellan SL (2014) Population dynamics and ecology of Arcobacter in sewage. Front Microbiol 5:525, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4224126&tool=pmcentrez&rendertype=abstract

Fleischer J, Schlafmann K, Otchwemah R, Botzenhart K (2000) Elimination of enteroviruses, other enteric viruses, F-specific coliphages, somatic coliphages and E. coli in four sewage treatment plants of southern Germany. J Water Supply Res Technol AQUA 49(3):127–38, Available from: http://aqua.iwaponline.com/content/49/3/127.abstract

Francy DS, Stelzer EA, Bushon RN, Brady AMG, Williston AG, Riddell KR et al (2012) Comparative effectiveness of membrane bioreactors, conventional secondary treatment, and chlorine and UV disinfection to remove microorganisms from municipal wastewaters. Water Res 46(13):4164–78, Available from: http://www.ncbi.nlm.nih.gov/pubmed/22682268

FSA (2006) Review of the analytical terminology for codex use in the procedural manual. FOOD Stand AGENCY Inf Bull METHODS Anal Sampl Foodst 68(March 2006):1–15

Garcia A, Goñi P, Cieloszyk J, Fernandez MT, Calvo-Beguería L, Rubio E et al (2013) Identification of free-living amoebae and amoeba-associated bacteria from reservoirs and water treatment plants by molecular techniques. Environ Sci Technol 47(7):3132–40

Gerba CP, Goyal SM, LaBelle RL, Cech I, Bodgan GF (1979) Failure of indicator bacteria to reflect the occurrence of enteroviruses in marine waters. Am J Public Health 69(11):1116–9, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1619291&tool=pmcentrez&rendertype=abstract

Goyal B (2013) Case study on evaluation of the performance of waste stabilization pond system of Jodhpur, Rajasthan, India. Water Pract Technol 8(1):95–104, Available from: http://wpt.iwaponline.com/content/8/1/wpt.2013011.abstract

Greub G, Raoult D (2004) Microorganisms resistant to free-living amoebae. Clin Microbiol Rev 17(2):413–33

Gurría A (2012) OECD environmental outlook to 2050: the consequences of inaction - key facts and figures [Internet]. Eur. World. [cited 2015 Nov 20]. Available from: http://www.oecd.org/env/indicators-modelling-outlooks/oecdenvironmentaloutlooktto2050theconsequencesofinaction-keyfactsandfigures.htm

Haramoto E, Katayama H, Oguma K, Yamashita H, Tajima A, Nakajima H et al (2006) Seasonal profiles of human noroviruses and indicator

bacteria in a wastewater treatment plant in Tokyo, Japan. Water Sci Technol 54(11-12):301–8, Available from: http://www.ncbi.nlm.nih.gov/pubmed/17302333

Harwood VJ, Levine AD, Scott TM, Chivukula V, Lukasik J, Farrah SR et al (2005) Validity of the indicator organism paradigm for pathogen reduction in reclaimed water and public health protection. Appl Environ Microbiol 71(6):3163–70, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1151840&tool=pmcentrez&rendertype=abstract

Hassani L, Imziln B, Boussaid A, Gauthier MJ (1992) Seasonal incidence of and antibiotic resistance among Aeromonas species isolated from domestic wastewater before and after treatment in stabilization ponds. Microb Ecol 23(3):227–37, Available from: http://www.ncbi.nlm.nih.gov/pubmed/24192933

Hernroth BE, Conden-Hansson A-C, Rehnstam-Holm A-S, et al (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

Herpers BL, de Jongh BM, van der Zwaluw K, van Hannen EJ (2003) Real-time PCR assay targets the 23S-5S spacer for direct detection and differentiation of Legionella spp. and Legionella pneumophila. J Clin Microbiol 41(10):4815–6, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=254366&tool=pmcentrez&rendertype=abstract

Hoekstra AY, Mekonnen MM, Chapagain AK, Mathews RE, Richter BD (2012) Global monthly water scarcity: blue water footprints versus blue water availability. Añel JA, editor. PLoS One 7(2):e32688, Available from: http://dx.plos.org/10.1371/journal.pone.0032688

Hot D, Legeay O, Jacques J, Gantzer C, Caudrelier Y, Guyard K et al (2003) Detection of somatic phages, infectious enteroviruses and enterovirus genomes as indicators of human enteric viral pollution in surface water. Water Res 37(19):4703–10, Available from: http://www.ncbi.nlm.nih.gov/pubmed/14568057

International Organization for Standardization (1998a) ISO 7899-1:1998 water quality—detection and enumeration of intestinal enterococci—part 1: miniaturized method (Most Probable Number) for surface and waste water

International Organization for Standardization (1998b) ISO 11731:1998. Water quality – detection and enumeration of Legionella

International Organization for Standardization (1999) ISO 6222:1999 water quality – enumeration of culturable micro-organisms – colony count by inoculation in a nutrient agar culture medium

International Organization for Standardization (2005) ISO 8199:2005. Water quality—general guidance on the enumeration of micro-organisms by culture

International Organization for Standardization (2012) ISO 9308-2:2012 water quality—enumeration of Escherichia coli and coliform bacteria—part 2: most probable number method

Jiang SC (2006) Human adenoviruses in water: occurrence and health implications: a critical review †. Environ Sci Technol 40(23):7132–40, Available from: http://dx.doi.org/10.1021/es060892o

Jjemba PK, Weinrich LA, Cheng W, Giraldo E, Lechevallier MW (2010) Regrowth of potential opportunistic pathogens and algae in reclaimed-water distribution systems. Appl Environ Microbiol 76(13):4169–78, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2897422&tool=pmcentrez&rendertype=abstract

Joly P, Falconnet P-A, André J, Weill N, Reyrolle M, Vandenesch F et al (2006) Quantitative real-time Legionella PCR for environmental water samples: data interpretation. Appl Environ Microbiol 72(4):2801–8, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1449029&tool=pmcentrez&rendertype=abstract

Jurzik L, Hamza IA, Wilhelm M (2015) Investigating the reduction of human adenovirus (HAdV) and human polyomavirus (HPyV) in a sewage treatment plant with a polishing pond as a tertiary treatment. Water Air Soil Pollut 226(9):284, Available from: http://link.springer.com/10.1007/s11270-015-2545-9

Kageyama T, Kojima S, Shinohara M, Uchida K, Fukushi S, Hoshino FB et al (2003) Broadly reactive and highly sensitive assay for Norwalk-like viruses based on real-time quantitative reverse transcription-PCR. J Clin Microbiol 41(4):1548–57, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=153860&tool=pmcentrez&rendertype=abstract

Katayama H, Haramoto E, Oguma K, Yamashita H, Tajima A, Nakajima H et al (2008) One-year monthly quantitative survey of noroviruses, enteroviruses, and adenoviruses in wastewater collected from six plants in Japan. Water Res 42(6-7):1441–8, Available from: http://www.ncbi.nlm.nih.gov/pubmed/17996920

Koo HL, Ajami N, Atmar RL, DuPont HL (2010) Noroviruses: the leading cause of gastroenteritis worldwide. Discov Med 10(50):61–70, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3150746&tool=pmcentrez&rendertype=abstract

Kozak GK, MacDonald D, Landry L, Farber JM (2013) Foodborne outbreaks in Canada linked to produce: 2001 through 2009. J Food Prot 76(1):173–83, Available from: http://www.ncbi.nlm.nih.gov/pubmed/23317873

Kroneman A, Verhoef L, Harris J, Vennema H, Duizer E, van Duynhoven Y et al (2008) Analysis of integrated virological and epidemiological reports of norovirus outbreaks collected within the Foodborne Viruses in Europe network from 1 July 2001 to 30 June 2006. J Clin Microbiol 46(9):2959–65, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2546741&tool=pmcentrez&rendertype=abstract

Legrand-Abravanel F, Mansuy J-M, Dubois M, Kamar N, Peron J-M, Rostaing L et al (2009) Hepatitis E virus genotype 3 diversity, France. Emerg Infect Dis 15(1):110–4, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2660688&tool=pmcentrez&rendertype=abstract

Lipp EK, Farrah SA, Rose JB (2001) Assessment and impact of microbial fecal pollution and human enteric pathogens in a coastal community. Mar Pollut Bull 42(4):286–93, Available from: http://www.ncbi.nlm.nih.gov/pubmed/11381749

Lodder WJ, de Roda Husman AM (2005) Presence of noroviruses and other enteric viruses in sewage and surface waters in The Netherlands. Appl Environ Microbiol 71(3):1453–61, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1065170&tool=pmcentrez&rendertype=abstract

Loisy F, Atmar RL, Guillon P, Le Cann P, Pommepuy M, Le Guyader FS (2005) Real-time RT-PCR for norovirus screening in shellfish. J Virol Methods 123(1):1–7, Available from: http://www.ncbi.nlm.nih.gov/pubmed/15582692

Lopman B, Vennema H, Kohli E, Pothier P, Sanchez A, Negredo A et al (2004) Increase in viral gastroenteritis outbreaks in Europe and epidemic spread of new norovirus variant. Lancet (Lond Engl) 363(9410):682–8, Available from: http://www.ncbi.nlm.nih.gov/pubmed/15001325

Lorenzo-Morales J, Coronado-Alvarez N, Martínez-Carretero E, Maciver SK, Valladares B (2007) Detection of four adenovirus serotypes within water-isolated strains of Acanthamoeba in the Canary Islands, Spain. Am J Trop Med Hyg 77(4):753–6, Available from: http://www.ajtmh.org/content/77/4/753.abstract

Magnet a, Fenoy S, Galván a L, Izquierdo F, Rueda C, Fernandez Vadillo C et al (2013) A year long study of the presence of free living amoeba in Spain. Water Res 47(19):6966–72

Mara DD, Pearson HW (1987) Waste stabilization ponds: design manual for mediterranean Europe [Internet]. [cited 2015 Jul 19]. Available from: https://books.google.es/books/about/Waste_Stabilization_Ponds.html?id=D_8nAQAAMAAJ&pgis=1

MARA DD, MILLS SW, PEARSON HW, ALABASTER GP (1992) Waste stabilization ponds: a viable alternative for small community

treatment systems. Water Environ J 6(3):72–8, Available from: http://doi.wiley.com/10.1111/j.1747-6593.1992.tb00740.x

Marzouk Y (1980) Relationship of viruses and indicator bacteria in water and wastewater of Israel. Water Res 14(11):1585–90, Available from: http://www.sciencedirect.com/science/article/pii/0043135480900627

Masclaux FG, Hotz P, Friedli D, Savova-Bianchi D, Oppliger A (2013) High occurrence of hepatitis E virus in samples from wastewater treatment plants in Switzerland and comparison with other enteric viruses. Water Res 47(14):5101–9, Available from: http://www.sciencedirect.com/science/article/pii/S0043135413004806

Maunula L, Kaupke A, Vasickova P, Söderberg K, Kozyra I, Lazic S et al (2013) Tracing enteric viruses in the European berry fruit supply chain. Int J Food Microbiol 167(2):177–85, Available from: http://www.ncbi.nlm.nih.gov/pubmed/24135674

Maynard HE, Ouki SK, Williams SC (1999) Tertiary lagoons: a review of removal mecnisms and performance. Water Res 33(1):1–13, Available from: http://www.ingentaconnect.com/content/els/00431354/1999/00000033/00000001/art00198

Medema G, Wullings B, Roeleveld P, van der Kooij D (2004) Risk assessment of Legionella and enteric pathogens in sewage treatment works. Water Sci Technol Water Supply [Internet]. [cited 2015 Nov 18];125–32. Available from: http://bvsalud.org/portal/resource/esSiqueira/rep-18894

Merga JY, Royden A, Pandey AK, Williams NJ (2014) Arcobacter spp. isolated from untreated domestic effluent. Lett Appl Microbiol 59(1):122–6, Available from: http://www.ncbi.nlm.nih.gov/pubmed/24666283

Monfort P, Baleux B (1990) Dynamics of Aeromonas hydrophila, Aeromonas sobria, and Aeromonas caviae in a sewage treatment pond. Appl Environ Microbiol 56(7):1999–2006, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=184551&tool=pmcentrez&rendertype=abstract

Niquette P, Servais P, Savoir R (2001) Bacterial Dynamics in the drinking water distribution system of Brussels. Water Res 35(3):675–82, Available from: http://www.sciencedirect.com/science/article/pii/S0043135400003031

Nordgren J, Matussek A, Mattsson A, Svensson L, Lindgren P-E (2009) Prevalence of norovirus and factors influencing virus concentrations during one year in a full-scale wastewater treatment plant. Water Res 43(4):1117–25, Available from: http://www.sciencedirect.com/science/article/pii/S0043135408005940

Ogorzaly L, Cauchie H-M, Penny C, Perrin A, Gantzer C, Bertrand I (2013) Two-day detection of infectious enteric and non-enteric adenoviruses by improved ICC-qPCR. Appl Microbiol Biotechnol 97(9):4159–66, Available from: http://www.ncbi.nlm.nih.gov/pubmed/23558583

Oki T, Kanae S (2006) Global hydrological cycles and world water resources. Science 313(5790):1068–72, Available from: http://www.sciencemag.org/content/313/5790/1068

Oragui JI, Curtis TP, Silva SA, Mara DD (2011) The removal of excreted bacteria and viruses in deep waste stabilization ponds in Northeast Brazil. IWA Publishing. Feb 16 [cited 2014 Dec 30]; Available from: http://www.iwaponline.com/wst/01810/wst018100031.htm

Ottoson J, Hansen A, Björlenius B, Norder H, Stenström TA (2006a) Removal of viruses, parasitic protozoa and microbial indicators in conventional and membrane processes in a wastewater pilot plant. Water Res 40(7):1449–57, Available from: http://www.sciencedirect.com/science/article/pii/S0043135406000789

Ottoson J, Hansen A, Björlenius B, Norder H, Stenström TA (2006b) Removal of viruses, parasitic protozoa and microbial indicators in conventional and membrane processes in a wastewater pilot plant. Water Res 40(7):1449–57, Available from: http://www.ncbi.nlm.nih.gov/pubmed/16533517

Pal A, Sirota L, Maudru T, et al (2006) Real-time, quantitative PCR assays for the detection of virus-specific DNA in samples with mixed populations of polyomaviruses. J Virol Methods 135:32–42. doi:10.1016/j.jviromet.2006.01.018

Palmer CJ, Tsai YL, Paszko-Kolva C, Mayer C, Sangermano LR (1993) Detection of Legionella species in sewage and ocean water by polymerase chain reaction, direct fluorescent-antibody, and plate culture methods. Appl Environ Microbiol 59(11):3618–24, Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=182507&tool=pmcentrez&rendertype=abstract

Pay Drechsel CAS, Liqa Raschid-Sally MR, Bahri A, Wastewater Irrigation and Health (2010) Assessing and mitigating risk in low-income countries. Dunstan house, London

Peña M, Rodriguéz J, Mara DD, Sepulveda M (2000) UASBs or anaerobic ponds in warm climates? a preliminary answer from Colombia. Water Sci Technol [Internet]. [cited 2015 Nov 18];59–65. Available from: http://www.bvsde.paho.org/bvsacd/leeds/uasb.pdf

Pusch D, Oh D-Y, Wolf S, Dumke R, Schröter-Bobsin U, Höhne M et al (2005) Detection of enteric viruses and bacterial indicators in German environmental waters. Arch Virol 150(5):929–47, Available from: http://www.ncbi.nlm.nih.gov/pubmed/15645371

Richards AM, Von Dwingelo JE, Price CT, Abu Kwaik Y (2013) Cellular microbiology and molecular ecology of Legionella-amoeba interaction. Virulence 4(4):307–14, Available from: http://www.tandfonline.com/doi/full/10.4161/viru.24290

Riera-Montes M, Brus Sjölander K, Allestam G, Hallin E, Hedlund K-O, Löfdahl M (2011) Waterborne norovirus outbreak in a municipal drinking-water supply in Sweden. Epidemiol Infect 139(12):1928–35, Available from: http://www.ncbi.nlm.nih.gov/pubmed/21251349

Rodríguez-Martínez S, Blanky M, Friedler E, Halpern M (2015) Legionella spp. isolation and quantification from greywater. MethodsX [Internet]. Nov [cited 2015 Nov 21]; Available from: http://www.sciencedirect.com/science/article/pii/S221501611500059X

Rusiñol M, Fernandez-Cassi X, Timoneda N, Carratalà A, Abril JF, Silvera C et al (2015) Evidence of viral dissemination and seasonality in a Mediterranean river catchment: implications for water pollution management. J Environ Manage 159:58–67, Available from: http://www.ncbi.nlm.nih.gov/pubmed/26046988

Sales-Ortells H, Fernandez-Cassi X, Timoneda N, Dürig W, Girones R, Medema G (2015) Health risks derived from consumption of lettuces irrigated with tertiary effluent containing norovirus. Food Res Int 68:70–7, Available from: http://www.sciencedirect.com/science/article/pii/S0963996914005560

Scheid P, Schwarzenberger R (2012) Acanthamoeba spp. as vehicle and reservoir of adenoviruses. Parasitol Res 111(1):479–85, Available from: http://www.ncbi.nlm.nih.gov/pubmed/22290448

Serrano-Suárez A, Araujo R (2013) Evaluation of Legionella presence in the water system of a public building by semi-nested polymerase chain reaction as a rapid screening method complementary to plate count. Water Sci Technol Water Supply 13(6):1560, Available from: http://www.iwaponline.com/ws/01306/ws013061560.htm

Serrano-Suárez A, Dellundé J, Salvadó H, Cervero-Aragó S, Méndez J, Canals O et al (2013) Microbial and physicochemical parameters associated with Legionella contamination in hot water recirculation systems. Environ Sci Pollut Res Int 20(8):5534–44, Available from: http://www.ncbi.nlm.nih.gov/pubmed/23436060

Shuval HI (1990) Summary—wastewater irrigation in developing countries—health effects and technical solutions. UNDP World Bank Water Sanit Progr 51:74

Sivapalasingam S, Friedman CR, Cohen L, Tauxe RV (2004) Fresh produce: a growing cause of outbreaks of foodborne illness in the United States, 1973 through 1997. J Food Prot 67(10):2342–53, Available from: http://www.ncbi.nlm.nih.gov/pubmed/15508656

Steinert M, Emödy L, Amann R, Hacker J (1997) Resuscitation of viable but nonculturable Legionella pneumophila Philadelphia JR32 by Acanthamoeba castellanii. Appl Environ Microbiol 63(5):2047–53

Teunis PFM, Moe CL, Liu P, Miller SE, Lindesmith L, Baric RS et al (2008) Norwalk virus: how infectious is it? J Med Virol 80(8):1468–76, Available from: http://www.ncbi.nlm.nih.gov/pubmed/18551613

Tyagi VK, Kazmi AA, Chopra AK (2008) Removal of fecal indicators and pathogens in a waste stabilization pond system treating municipal wastewater in India. Water Environ Res 80(11):2111–7, Available from: http://www.ingentaconnect.com/content/wef/wer/2008/00000080/00000011/art00004

Wheaton TA, Castle WS (2001) High application rates of reclaimed water benefit citrus tree growth and fruit production. HortSci 36:1273–7, Available from: http://www.mendeley.com/catalog/high-application-rates-reclaimed-water-benefit-citrus-tree-growth-fruit-production/

ARTICLE II

Avaluació dels mètodes de concentració i extracció de virus a partir d'aigua residual en el context de seqüenciació de metagenomes vírics

2. Avaluació dels mètodes de concentració i extracció de virus a partir d'aigua residual en el context de seqüenciació de metagenomes vírics

L'estudi del viroma present a l'aigua residual urbana constitueix una eina interessant per a la vigilància i el desenvolupament d'estudis epidemiològics, així com per a avaluar l'eficiència dels tractaments aplicats a les plantes de tractament d'aigües residuals urbanes.

Per estudiar els virus a l'ambient cal aplicar processos de concentració que permetin separar els virus d'altres components, a la vegada que aconseguim retenir-los en un menor volum. Posteriorment a la seva concentració cal aplicar un procés d'extracció per aïllar-ne el material genètic que pot ser de característiques molt diferents: cadena senzilla o cadena doble, genoma linear o circular, segmentat o no segmentat, ADN o ARN o fins i tot, combinacions d'ambdós.

L'estudi presentat en aquest capítol analitza l'efecte que tenen els mètodes de concentració i extracció en la composició del viroma aplicant tècniques de seqüenciació massiva. Amb aquesta finalitat, es va mostrejar un gran volum d'aigua residual que va ser dopat amb dos virus: HAdV35 i Norovirus murí. Es van avaluar 4 mètodes de concentració diferents, tal i com apareixen descrits a la literatura, incloent la floculació amb llet descremada (SMF), la precipitació amb PEG, l'ús de columnes MAF o la llana de vidre. Simultàniament es van testar 4 kits d'extracció de ADN/ARN: Nucleospin RNA XS, QIAmp viral RNA mini kit, NucliSENS® miniMAG® o PowerViral® Environmental ARN/ADN isolation kits. En total es van assajar 16 combinacions de mètodes de concentració/extracció per triplicat i un control negatiu per a cadascuna de les combinacions que van ser analitzades per tècniques de seqüenciació massiva. Es van utilitzar diferents paràmetres per avaluar l'efecte sobre la composició viral: com la *viral richness,* percentatge de *reads* associats a virus o detecció de virus patògens. A més les

mostres van ser analitzades en paral·lel mitjançant q(RT)PCR específiques pels dos virus afegits a la mostra.

En total 14 famílies víriques que inclouen patògens humans van ser detectades en aquest estudi. Els resultats senyalen que els mètodes de concentració i extracció exerceixen un efecte sobre les poblacions virals detectades aplicant tècniques de seqüenciació massiva. La combinació que va permetre detectar un major nombre d'espècies virals diferents van ser les que utilitzaven QIAmp viral RNA mini kit o PowerViral® Environmental ARN/ADN isolation kit. El major nombre de *reads* associats a virus va ser utilitzant concentració amb PEG i extracció amb Nucleospin RNA XS. També s'ha vist que existeix una correlació important entre els *reads* per milió (RPM) obtinguts per HAdV i la seva quantificació en còpies genòmiques per qPCR. Aquesta correlació no s'observa per a norovirus murí.

PLOS ONE

RESEARCH ARTICLE

Evaluation of Methods for the Concentration and Extraction of Viruses from Sewage in the Context of Metagenomic Sequencing

Mathis Hjort Hjelmsø[1☯*], Maria Hellmér[2☯], Xavier Fernandez-Cassi[3], Natàlia Timoneda[3,4], Oksana Lukjancenko[1], Michael Seidel[5], Dennis Elsässer[5], Frank M. Aarestrup[1], Charlotta Löfström[2¤], Sílvia Bofill-Mas[3], Josep F. Abril[3,4], Rosina Girones[3], Anna Charlotte Schultz[2]

1 Research Group for Genomic Epidemiology, The National Food Institute, Technical University of Denmark, Kongens Lyngby, Denmark, **2** Division of Microbiology and Production, The National Food Institute, Technical University of Denmark, Søborg, Denmark, **3** Laboratory of Virus Contaminants of Water and Food, Department of Genetics, Microbiology, and Statistics, University of Barcelona, Barcelona, Catalonia, Spain, **4** Institute of Biomedicine of the University of Barcelona, University of Barcelona, Barcelona, Catalonia, Spain, **5** Institute of Hydrochemistry, Chair of Analytical Chemistry, Technical University of Munich, Munich, Germany

☯ These authors contributed equally to this work.
¤ Current address: Food and Bioscience, SP Technical Research Institute of Sweden, Lund, Sweden
* mhjhj@food.dtu.dk

Citation: Hjelmsø MH, Hellmér M, Fernandez-Cassi X, Timoneda N, Lukjancenko O, Seidel M, et al. (2017) Evaluation of Methods for the Concentration and Extraction of Viruses from Sewage in the Context of Metagenomic Sequencing. PLoS ONE 12(1): e0170199. doi:10.1371/journal.pone.0170199

Editor: Patrick Tang, Sidra Medical and Research Center, QATAR

Received: October 13, 2016

Accepted: January 2, 2017

Published: January 18, 2017

Data Availability Statement: The metagenomic sequences are available from the European Nucleotide Archive (ENA) at the European Bioinformatics Institute (EBI) under accession number PRJEB15242.

Funding: Funding and support has been provided by European Union's Horizon 2020 research and innovation programme, COMPARE, grant agreement No. 643476 (http://www.compare-europe.eu/), the JPI Water project, METAWATER

Abstract

Viral sewage metagenomics is a novel field of study used for surveillance, epidemiological studies, and evaluation of waste water treatment efficiency. In raw sewage human waste is mixed with household, industrial and drainage water, and virus particles are, therefore, only found in low concentrations. This necessitates a step of sample concentration to allow for sensitive virus detection. Additionally, viruses harbor a large diversity of both surface and genome structures, which makes universal viral genomic extraction difficult. Current studies have tackled these challenges in many different ways employing a wide range of viral concentration and extraction procedures. However, there is limited knowledge of the efficacy and inherent biases associated with these methods in respect to viral sewage metagenomics, hampering the development of this field. By the use of next generation sequencing this study aimed to evaluate the efficiency of four commonly applied viral concentrations techniques (precipitation with polyethylene glycol, organic flocculation with skim milk, monolithic adsorption filtration and glass wool filtration) and extraction methods (Nucleospin RNA XS, QIAamp Viral RNA Mini Kit, NucliSENS® miniMAG®, or PowerViral® Environmental RNA/DNA Isolation Kit) to determine the viriome in a sewage sample. We found a significant influence of concentration and extraction protocols on the detected viriome. The viral richness was largest in samples extracted with QIAamp Viral RNA Mini Kit or PowerViral® Environmental RNA/DNA Isolation Kit. Highest viral specificity were found in samples concentrated by precipitation with polyethylene glycol or extracted with Nucleospin RNA XS. Detection of viral pathogens depended on the method used. These results contribute to the understanding of method associated biases, within the field of

(http://www.waterjpi.eu/images/Kick-Off/METAWATER.pdf) and Aquavalens (EU FP7-KBBE-2012-6) (http://aquavalens.org/). This study was partially funded by a grant of the Catalan Government as the Consolidated Research Group VirBaP (2014SRG914) (http://www.ub.edu/microbiologia/grupbacterisen/index.html). During the development of the study XFC was a fellow of the Catalan Government "AGAUR" (FI-DGR) (http://www.uab.cat/web/research/uab-research-training-grants/postdoctoral-grants/catalan-government-1184220108167.html) and NT was a fellow of the Spanish Ministry of Science (http://www.uab.cat/web/research/uab-research-training-grants/postdoctoral-grants/ministry-of-education-and-science-1184220108300.html). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

viral sewage metagenomics, making evaluation of the current literature easier and helping with the design of future studies.

Introduction

Within raw sewage, feces, urine and other biological fluids from thousands of humans are mixed together with food and household waste, industrial waste, and runoff water. Every individual, who is connected to the drainage system, contributes with his or hers own microbiota [1], including infecting pathogens [2]. This makes sewage an attractive matrix for epidemiological studies [3], microbial source tracking [4], and for controlling the efficacy of pathogen removal in wastewater treatment plants [5,6]. Sewage has been shown to harbor a diverse viral population including enteric, respiratory and oncogenic viruses [7]. The high viral diversity and the continuous mutation of viral species makes identification with traditional methods difficult and time consuming, therefore many studies have turned to Next Generation Sequencing (NGS) approaches instead [7–9]. Metagenomic sequencing of the virus associated nucleic acids is considered to be an unbiased approach enabling the detection of all known viral species, as well as the discovery of novel and emergent species [10]. Three main challenges exist for viral sewage metagenomics. First, only a small fraction of the total nucleic acids are of known viral origin, hence mechanical and enzymatic viral purification is often needed [9]. Second, the low abundance of viral particles in the samples requires the use of viral concentration methods prior to nucleic acid extraction [11] and is often combined with subsequent random DNA amplification [12]. Third, the nucleic acid extraction procedure has to cover the large variety in viral structures and genome types. To overcome these biases, different methods to concentrate viruses from water samples have been developed, including: polyethylene glycol precipitation (PEG) [8], $FeCl_3$ precipitation [13], skimmed milk flocculation (SMF) [14], glass wool filtration (GW) [15] or monolithic adsorption filtration (MAF) [16]. The influence of concentration method on viral recovery has been evaluated on sea water [17], spiked tap water [15,18] and raw sewage [19], cautioning of method associated biases. To our knowledge, no major comparison studies using metagenomics have been performed with sewage water.

Biases caused by nucleic acid extraction kits have been well documented for both bacteria [20,21] and viruses [22,23]. In addition, contaminants have been found to be ubiquitous in some extraction kits [24] and laboratory reagents [25], potentially giving rise to false positive results [26,27]. A better understanding of specific method associated biases, in respect to viral wastewater metagenomics, would make evaluation of the current literature easier, and help guide future studies.

In this study we evaluated four previously published concentration methods, PEG, MAF, SMF, and GW, as well as four extraction kits, Nucleospin RNA XS (NUC), QIAamp Viral RNA Mini Kit (QIA), NucliSENS® miniMAG® (MIN), or PowerViral® Environmental RNA/DNA Isolation Kit (POW), for wastewater viral metagenomics, in a full factorial design resulting in 16 combinations of procedures. Aspects studied included viral community composition, viral selectiveness, viral richness, viral pathogen detection, and viral contaminants. Extracted nucleotides were amplified with PCR and sequenced using the Illumina MiSeq platform.

Materials and Methods

Sample collection, spiking and pooling

In July 2015 raw sewage (130 L) was collected at the waste water treatment plant BIOFOS Lynetten in Copenhagen, Denmark, receiving waste water from about 550,000 inhabitants. Approval was granted from BIOFOS Lynettefællesskabet A/S before sampling. The sewage was mixed thoroughly in a single container and spiked to a concentration of 1.74×10^8 RT-PCR units/L of murine norovirus (MNV) (kindly provided by Dr Virgin, Washington University School of Medicine, USA), and 2.13×10^9 genome copies/L of human adenovirus 35 (HAdV). The sample was mixed for 5 min before aliquoted and stored at—20˚C until further processing.

Concentration methods

Four different methods were used to concentrate virions from the sewage samples: protein precipitation with PEG, organic flocculation with SMF and filtration with positively charged filters, MAF, or GW. All concentration methods were done in triplicate together with a negative control using sterile molecular grade water (VWR—Bie & Berntsen, Søborg, Denmark).

PEG

The PEG protocol was based on the procedure as previously described [8]. Initially, 25 mL of glycin buffer (0.05 M glycine, 3% beef extract, pH 9.6) was added to 200 mL of sewage and mixed, to detach virions bound to organic material. The sample was then centrifuged at 8,000×g for 30 min, and the collected supernatant was filtered through a 0.45 μm polyethersulfone (PES) membrane (Jet Biofil, Guangzhou, China) to remove bacterial and eukaryotic cells. Viruses were precipitated from the supernatant by incubation with PEG 8000 (80 g/L) and NaCl (17.5 g/L) during agitation (100 rpm) overnight at 4˚C, followed by centrifugation for 90 min at 13,000×g. The resulting viral-containing pellet was eluted in 1 mL phosphate buffer saline (PBS) and stored at -80˚C until further processing.

MAF

The principle of the MAF adsorption/elution method was based on the procedure as previously described [18]. Monolithic discs, diameter 3.86 cm and height 1.0 cm, were synthesized by polymerization of polyglycerol-3-glycidyl ether (Ipox chemicals, Laupheim, Germany). An 80:20 mixture of toluene and tert-buthyl methyl ether was used as porogen to create monoliths with a pore size of ca. 20 μm. After synthesis, functionalization was performed by recirculating 10% diethylamine in 50% ethanol at 60˚C through the monolithic disks for 3 h to create positively charged diethylaminoethyl groups on the pore surface. Afterwards the monoliths were rinsed with ultrapure water and stored at 4˚C until further use. One liter of raw sewage was filtrated through a MAF disc (Microarray and Bioseparation Group of the Institute of Hydrochemistry, Technical University of Munich, Germany) assembled as previously described [28]. Viruses were eluted from the filter by soaking 2×2 min in a total of 20 mL high salt buffer (1.5 M NaCl, 0.05 M HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) buffer, pH 7). The eluate was further concentrated to 3 mL by 100 kDa Amicon ultra centrifugation filters (Merck Millipore, Cork, Ireland) according to the manufacturer's instructions. The viral concentrate was stored at -80˚C until further processing.

SMF

Organic flocculation with skimmed milk was based on the procedure as previously described [7]. In brief, 100 mL pre flocculated skimmed milk solution (1% (w/v) skimmed milk powder

(Difco, Detroit, MI, USA), 3.2% (w/w) sea salts (Sigma Aldrich Chemie GMBH, Steinheim, Germany)) at pH 3.5 was added to 10 L of acidified (using HCl to pH 3.5) raw sewage and mixed for 8 h. Flocculants were allowed to sediment for 8 h, and centrifuged at 8,000×g for 40 min. The pelleted viral concentrate was suspended in 15 mL phosphate buffer (1:2 (v/v) mixture of 0.2 M Na_2HPO_4 and 0.2 M NaH_2PO_4). The phosphate suspension was eluted in 30 mL 0.25 M glycine buffer (pH 9.5) with slow agitation for 45 min at 4˚C. Suspended solids were separated by centrifugation at 8,000×g for 40 min at 4˚C. The sample was neutralized to pH 7 by adding 1 M HCl. Virions present in the supernatant were concentrated by ultracentrifugation at 90,000×g for 90 min (Sorvall Discovery 90SE) at 4˚C and suspended in 2 mL PBS. The viral concentrate was stored at -80˚C until further processing.

GW

The glass wool filters were prepared as previously described [29]. Sodocalcic glass wool (15 g) (Ouest Isol, Alizay, France) was packed into a PVC tube with the density of 0.11 g/cm^3 and pretreated with the following solutions, 100 mL NaOH (1 M) for 15 min, 1 L sterile distilled water, 100 mL HCl (1 M) for 15 min, and 1 L sterile distilled water. Samples of raw sewage (4 L) were filtered through the glass wool column. Viruses were eluted by incubating 100 mL elution buffer (3% beef extract, 0.5 M glycine, pH 9.5) for 15 min. Secondary concentration was done by PEG precipitation (as above) to a final volume of 1 mL.

DNase/RNase treatment + chloroform-butanol treatment

All viral concentrates were treated with OmniCleave endonuclease (Epicentre, Wisconsin, USA) to remove extracellular DNA/RNA as previously described [30]. Samples were further purified by extraction using a 1:1 mixture of chloroform-butanol [31] to remove nucleases and inhibitors.

Extraction methods

Nucleic acids were extracted from 200 μL-portions of the respective viral concentrate using four different extraction kits; NUC (Macherey-Nagel, Düren, Germany), QIA (Qiagen, Valencia CA, USA), MIN (BioMerieux, Herlev, Denmark) or POW (MO BIO, Carlsbad, CA, USA). In all cases, extractions were carried out according to manufacturer's instructions.

qPCR analysis of spiked viruses

Detection of HAdV and MNV was performed on extracted nucleic acids (undiluted and 10-fold diluted) in a 96-well plate format of ABI Step One (Applied Biosystems, Naerum, Denmark). MNV RNA was detected by quantitative reverse transcriptase polymerase reaction (qRT-PCR) using the RNA UltraSense one-step qRT-PCR system (Invitrogen, Taastrup, Denmark) and previously described primers and probes [32]. Amplification was performed in a 25 μL reaction mixture containing 5 μL extracted nucleic acids and 20 μL qRT-PCR reaction mixture with 500 nM forward primer, 900 nM reverse primer, 250 nM probe, 1 × UltraSense reaction mix, 1 × ROX reference dye and 1 × UltraSense enzyme mix under the following reaction conditions, 55˚C for 1 min and 95˚C for 5 min followed by 45 cycles of 95˚C for 15 s, 60˚C for 1 min, and 65˚C for 1 min. HAdV DNA was detected by qPCR using TaqMan Universal Master Mix (Applied Biosystems, Naerum, Denmark), and previously described primers and probe [33]. Amplification was performed in a total of 25 μL reaction mixture containing 5 μL extracted nucleic acids and 20 μL qPCR reaction mixture containing 1 × TaqMan Universal Master Mix, primer and probe concentrations and qPCR running

conditions are described in [33]. Quantification was performed using standard curves generated from 10-fold dilution series, of extracted RNA of cell propagated MNV or of ds HAdV DNA segments, artificially constructed by gBlocks® Gene Fragments (Integrated DNA Technologies, Leuven, Belgium).

Reverse transcriptase, library preparation and, sequencing

To prepare extracted RNA and DNA for sequencing, each viral extract was subjected to reverse transcriptase and PCR amplified, as previously described [34]. Briefly, first strand cDNA synthesis were performed using the SuperScript® III First-Strand Synthesis SuperMix (Invitrogen, Carlsbad, California) and 1 μL Primer A (50 μM) (5′-GTTTCCCAGTCAC GATCNNNNNNNNN-3′) according to the manufacturer's instructions. Second strand DNA synthesis were performed using Klenow Fragment exo-polymerase (Thermo Fisher Scientific, Waltham, MA, USA) as previously described [30]. Double stranded DNA products were PCR amplified using AmpliTaq Gold (Qiagen, Valencia CA, USA) as per manufacturer's instruction using 0.8 μM Primer B (5′- GTTTCCCAGTCACGATC -3′) and the following conditions, 10 min at 95˚C, 25 cycles of amplification (94˚C for 30 s, 40˚C for 30 s, 50˚C for 30 s and 72˚C for 1 min), and 1 cycle of elongation (72˚C for 10 min). PCR products were purified using the DNA Clean & Concentrator™-5 (Zymo Research, Irvine CA, USA). NGS library preparation was performed using the Nextera XT DNA Library Preparation kit (Illumina, Eidenhoven, The Netherlands) according to the manufacturer's instructions. The 64 samples were sequenced on three Illumina MiSeq runs with an average output of 1.4×10^6 250 bp paired-end reads per sample (S1 Table).

Bioinformatic analyses

The distribution of viral species was determined using MGmapper software version 2.2 (https://cge.cbs.dtu.dk/services/MGmapper/) [31]. The MGmapper tool follows three main steps: quality assessment of the raw reads, mapping of reads to the reference databases, and post-processing of mapping results. Quality assessment was done using cutadapt [35] which performs common adapter removal, trimming of the low-quality ends from reads with a minimum Phred quality score of 20, and later discards reads that are shorter than 40 bp. Later, already trimmed pair-end reads were aligned to a pre-defined set of reference sequence databases using bwa mem [36] ver. 0.7.7-r441 with default settings. In this study, reads were mapped against three viral reference databases (S2 Table): whole genomes virus sequences (Virus) and viral sequences extracted from nt database (Virus_nt), obtained from Genbank (http://www.ncbi.nlm.nih.gov/genbank/), as well as Vipr database (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3245011). Samtools [37] were used to remove singletons and filter reads where neither a read nor its mate is mapped. Reads were mapped in best-mode, meaning that mapping was performed against all databases, simultaneously, and later for each read pair the best hit among all alignments is chosen. A pair of reads is considered as a hit only if the sum of the alignment scores (SAS) is higher than any SAS values from other database hits. If a pair of reads has identical SAS values when mapping to several databases, the only one pair, associated with the database that was specified first in the list of reference databases, is kept. In the last, post-processing step, alignments are filtered based on matches/mis-matches threshold. In this analysis, 70% matches/mis-matches threshold needed to be satisfied in order the hit to be considered significant. The metagenomic sequences are available from the European Nucleotide Archive (ENA) at the European Bioinformatics Institute (EBI) under accession number PRJEB15242.

Statistics and plots

Viral richness was estimated using the program CatchAll [38] and the non-parametric Chao1 richness index, as a measure of number of viral species in a sample. All statistics were done in R [39], using two-way analysis of variance (ANOVA) test for determining the overall significance of concentration or extraction method on the studied factors (viral richness, etc.). Subsequently, pairwise t-tests with "Holm-Bonferroni" [40] p-value adjustments were applied to determine significant pairwise effects between individual concentration or extraction methods. Principal component analysis (PCA) were performed using prcomp and plotted with ggbiplot [41]. Heatmaps were created using pheatmap [42]. Linear regression between reads per million (RPM) and genome copies per liter were done in Excel on log transformed data.

Results

In this study different virus concentration and nucleic acid extraction methods were evaluated for metagenomic analysis of sewage samples. Sequencing results showed that the majority of the mapped reads (>80%) were of viral origin (S1 Fig). However, between 60 and 90% of the total reads were unmapped. The three main viral families detected were Adenoviridae (human viruses) including the spiked HAdV 35, Virgaviridae (plant viruses), and Siphoviridae (bacteriophages). The sequencing data were analyzed further to determine the viral community composition, viral specificity, viral richness, and detection of pathogenic species.

Viral community composition

To compare the viral community composition resulting from the individual concentration and extraction methods, a series of PCAs were made using the relative abundances from the nine most abundant viral families, accounting for more than 99% of the mapped viral sequences. The effect of extraction (Fig 1) and concentration (Fig 2) were plotted independently for easier visualization. Samples plotted close together have similar viral community compositions, whereas samples far away from each other are less alike. The negative controls clustered together far away from the samples in initial PCA plots (data not shown). To allow for better visualization of the effect of concentration and extraction on the sewage samples, they were not included in Figs 1 and 2.

Sample replicates extracted with NUC clustered away from the samples extracted with the other methods, when concentrated with PEG (Fig 1D). This was also true for the concentrates from MAF (Fig 1A), SMF (Fig 1B), and GW (Fig 1C), suggesting that the viral community composition of the NUC extractions differed from the other tested extraction methods. The samples extracted with POW and MIN clustered together, suggesting similar viral community compositions (Fig 1A–1D). The samples extracted with QIA sometimes clustered separately (Fig 1A and 1C) and sometimes together with the samples extracted with POW and MIN (Fig 1B and 1D). The four concentration methods formed separate non-overlapping clusters regardless of extraction kit used (Fig 2A–2D), although some variation between replicates were observed.

Viral specificity

The proportion of reads mapping to viruses ranged between 3.4% and 49.4%. Both the concentration and the extraction methods had a statistical significant effect on the viral specificity (two-way ANOVA, $p < 0.001$). However, a significant interacting effect (two-way ANOVA, $p < 0.001$) indicated that the effect on viral specificity by the extraction method was affected by the type of concentration method, and vice versa.

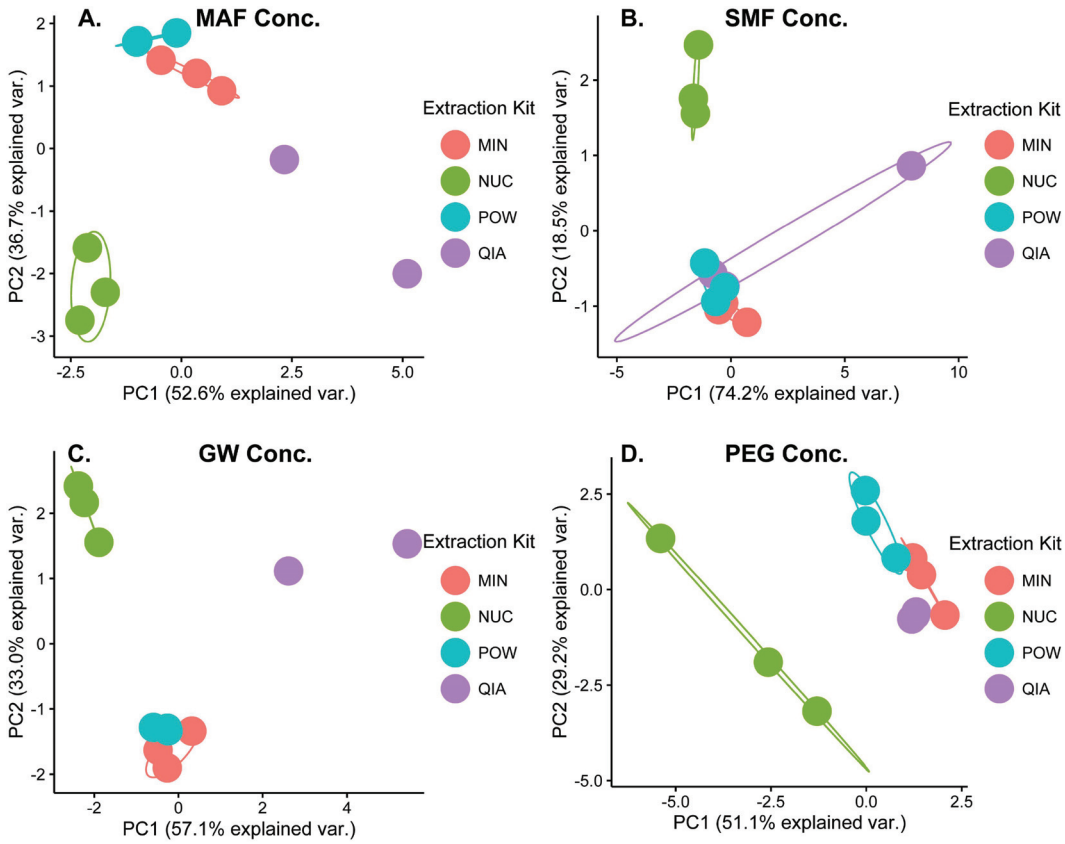


Fig 1. The influence of extraction method on the viral community composition. PCA plots made by using the relative abundances of the nine most abundant viral families. Separate PCAs were done for (A) samples concentrated with MAF, (B) SMF, (C) GW, and (D) PEG. Sample replicates were individually plotted and grouped according to the extraction method. In cases where only two samples were present, no ellipse representing the cluster was drawn.

doi:10.1371/journal.pone.0170199.g001

The PEG concentration method had a significant larger mean proportion of viral reads compared to the SMF and GW methods (pairwise t-test, $p < 0.01$) (Fig 3A). For the extraction methods, NUC had a significant larger mean proportion of viral reads compared to POW and QIA (pairwise t-test, $p < 0.01$) (Fig 3B). However, there were some interacting effects, with MIN scoring higher than NUC when used in combination with PEG and GW, implying that the MIN method depends heavily on the performance of the concentration method.

Viral richness

Both concentration and extraction methods had an effect on the viral richness. However, none of the concentration methods were statistically different from each other (pair-wise t-test) (Fig 4A). For the extraction methods, NUC had a significantly lower Chao1 richness than the other methods (Fig 4B). POW and QIA had the highest mean richness estimates of 516 and 495, respectively.

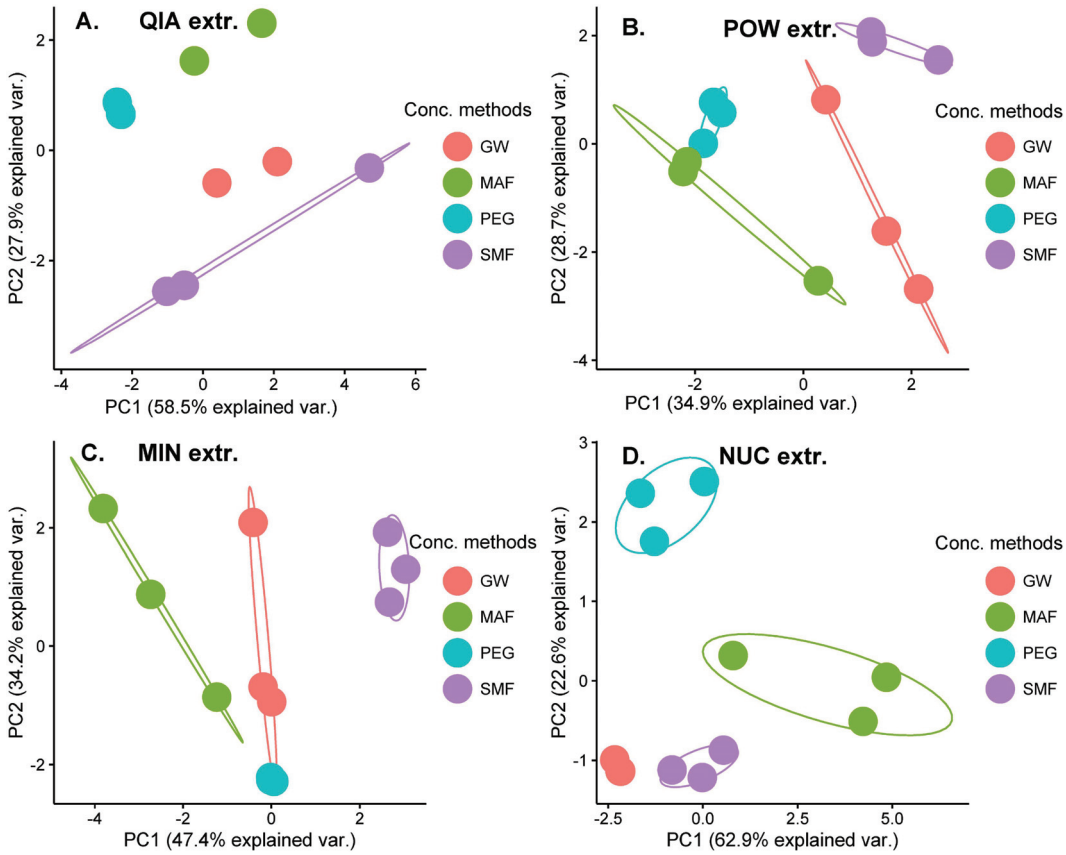


Fig 2. The influence of concentration method on the viral community composition. PCA plots made by using the relative abundances of the nine most abundant viral families. Separate PCAs were done for (A) samples extracted with QIA, (B) POW, (C) MIN, and (D) NUC. Sample replicates were individually plotted and grouped according to the concentration method. In cases where only two samples were present, no ellipse representing the cluster was drawn.

doi:10.1371/journal.pone.0170199.g002

Detection of pathogenic species

Fourteen viral families with suspected human pathogens were detected (Fig 5). The most prevalent was Adenoviridae including the spiked HAdV. The highest read count for the viral RNA families, Reoviridae, Picornaviridae, Astroviridae, Caliciviridae and Picorbinaviridae, was obtained in samples extracted with NUC. The spiked HAdV was detected at the highest abundance when extracted with MIN. The effect of the concentration methods was not as pronounced as for the extraction kits. The highest read count of the DNA virus family, Adenoviridae, was found in samples concentrated with MAF and PEG. In general, SMF had a lower performance compared with the other methods when testing for Adenoviruses such as the spiked HAdV. However, the combination of SMF and NUC had the highest read count for most of the RNA viruses.

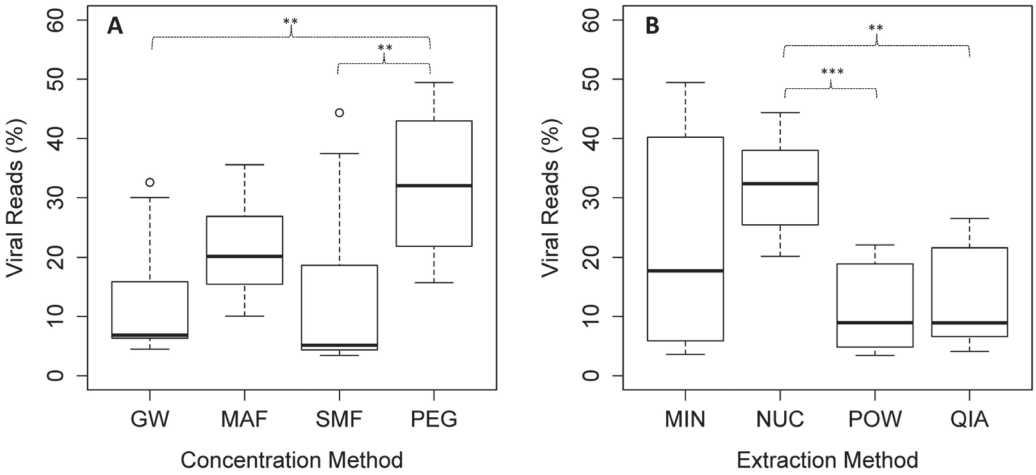


Fig 3. Viral selectivity measured in percentage of reads. (A) Viral selectivity for the tested concentration methods (B) and extraction methods. Each boxplot was made from 12 individual samples (including the four extraction/concentration methods with three replicates each). The bar, box, whiskers and circles represents median, inter-quartile range, inter-quartile range times 1.5, and outliers, respectively. Asterisks represent significance level of a pairwise t-test with "Holm-Bonferroni" adjusted p-values. ** = $p < 0.01$, *** = $p < 0.001$.

doi:10.1371/journal.pone.0170199.g003

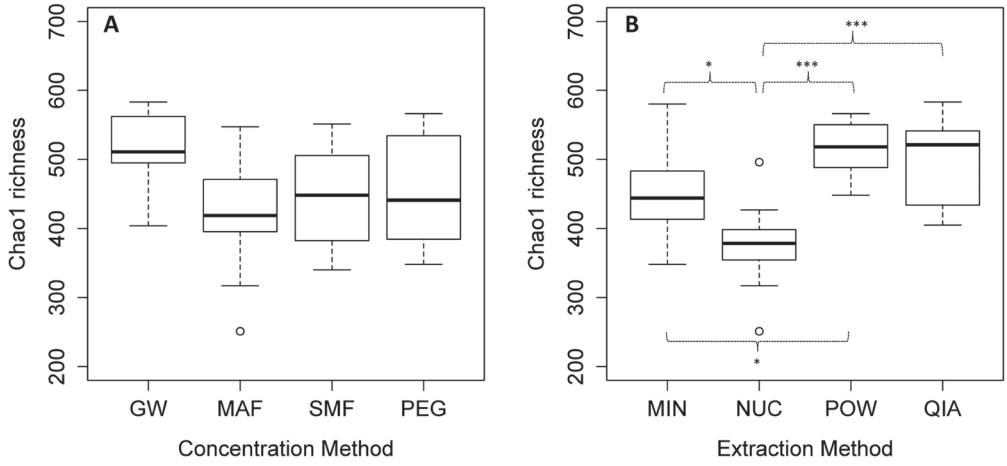


Fig 4. Viral species richness. (A) Viral Chao 1 species richness of the tested concentration methods, and (B) extraction methods. Each boxplot was made from 12 individual samples (including the four extraction/concentration methods with three replicates each). The bar, box, whiskers and circles represents median, inter-quartile range, inter-quartile range times 1.5, and outliers, respectively. Asterisks represent significance level of a pairwise t-test with "Holm-Bonferroni" adjusted p-values. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

doi:10.1371/journal.pone.0170199.g004

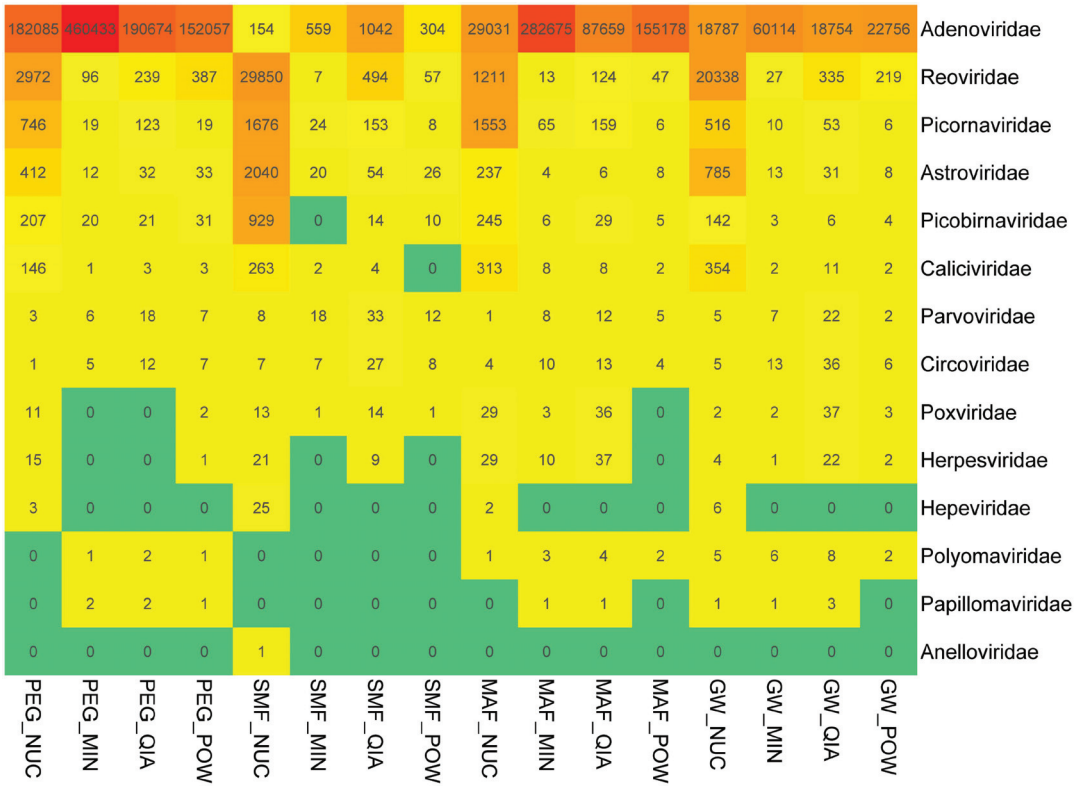


Fig 5. Detection of pathogenic viral families. Heatmap of the relative abundance of 14 human pathogenic viral families, detected by the 16 different concentration/extraction combinations. The numbers within each cell represents reads per million. The colours range from green = no detection, to red = high relative abundance.

doi:10.1371/journal.pone.0170199.g005

The spiked MNV were only detected by metagenomics in 57% of the samples, and at low read counts, from 2 to 194 reads. The combinations that could detect most MNV were MAF and GW extracted with NUC as well as MAF extracted with QIA and MIN.

qPCR analysis of spiked viruses

The detected concentrations of HAdV and MNV varied widely between the different method combinations (S2 and S3 Figs) with mean values ranging between 650 and 8.2×10^7 genome copies/L for HAdV, and 1.8×10^2 and 3.9×10^5 RT-PCR units/L for MNV. Choice of extraction method did not influence HAdV or MNV recovery. However, concentration methods had a significant impact (pairwise t-test, $p < 0.05$). The highest recovery of HAdV and MNV was obtained with PEG followed by MAF, GW, and SMF.

Inhibition

To investigate the possibility of PCR inhibition, extracts of nucleic acids (undiluted and 10-fold diluted) from all samples were analyzed for the two spiked viruses, MNV and HAdV,

with qPCR (S3 Table). The lowest inhibition of MNV and HAdV were observed in samples concentrated with PEG or extracted with MIN. Strongest inhibition was observed in samples concentrated with SMF for both MNV and HAdV. In addition samples extracted with QIA showed strong inhibition of HAdV detection.

Correlation between qPCR quantification and reads per million (RPM)

To investigate the correlation between viral concentrations and RPM, qPCR data was compared with read counts from the two spiked viruses, HAdV and MNV. There was a strong correlation between RPM and qPCR enumeration for HAdV ($R^2 = 0.82$). However, no relationship was observed for MNV ($R^2 = 0.07$).

Contamination

To detect method dependent contamination, a negative control was included, using sterile molecular grade water, for each of the 16 method combinations. Negative controls generally had a low total read count, a low percentage of viral reads (0.05–3.4%), and a high abundance of reads with human, bacterial, fungal and parasitic origin (S1 Fig). Several viral species were found in the negative controls with much higher RPM values than in the corresponding samples, suggesting that they originated from the corresponding kits or reagents. Reads mapping to pandora viruses, tupaiid herpes viruses, and *Citrobacter* phages were contaminants in all procedures except the ones using QIA extractions. However several mardi viruses were found exclusively in the QIA negative controls.

Discussion

In the presented study we evaluated the influence of four commonly applied concentration and extraction methods on viral metagenome analysis.

The viral community composition was heavily biased by the type of concentration procedure, which dramatically skewed the relative abundances (Fig 2). Choice of extraction kit did not influence the viral community composition to the same degree (Fig 1). However, the results from the NUC extraction kit were remarkably different from samples extracted with the three other kits. The NUC kit includes an "on column DNase step" after viral capsid disruption, which selects for RNA viruses and could explain the separate clustering in the PCA plots (Fig 1). Based on the results from this study it seems inadvisable to compare results, in relation to viral community composition, between studies using different concentration methods and to some degree also extraction methods.

A high species richness have been linked to several ecosystem functions [43], and is often included as a factor in ecological studies. In this study we included the measure to discern if some methods were better at catching the entire spectrum of viral species. Our results show that the choice of extraction method is of more importance than the choice of concentration method with regard to viral richness. However, samples concentrated with GW had a slightly higher richness compared to the other concentration methods (Fig 4A). The low mean richness of the samples extracted with NUC can probably be explained by the DNase step, degrading the genomes of DNA viruses and the species rich bacteriophages [44].

Viral specificity, or how large a fraction of the sequencing reads is of viral origin, is important for sensitivity reasons, increasing the chance to detect rare or less abundant species. A high viral specificity also has financial implications, causing large savings on both sequencing, and for subsequent CPU hours used in the bioinformatics analyses. In this study, the PEG protocol was the best concentration method, in respect to viral specificity (Fig 3A). This might be explained by the initial filtration step, not part of the other evaluated protocols. Pre-filtration

might have improved the viral specificity in the other concentration methods, although clotting might become a problem due to the increased volumes processed with these methods. The NUC had a consistent high viral specificity (three times that of POW and QIA), probably due to the effective removal of DNA from other organisms, and contaminants, during the DNase step. Overall, there was a 10-fold difference in viral specificity between the lowest and the highest method combination, highlighting the potential savings associated with choice of method. We observed a generally high viral specificity in this study compared to previous studies [45]. This might be due to the addition of the spiked HAdV, inflating the amount of virus in the sewage matrix, but should not have any influence on the method comparisons.

Sewage metagenomics is often used to detect human viral pathogens [8] including the important enteric RNA viruses as norovirus [46], rotavirus [47] and Hepatitis A and E virus [48] that has a big impact on public health [49]. These RNA viral families were best detected when using the NUC extraction kit compared to the other tested extraction kits, especially in combination with the concentration method SMF. However, if looking at DNA viruses exclusively, the MIN extraction combined with PEG, MAF, or GW may be preferable, since it produced the highest read counts for the spiked Adenoviridae. Low detection of Adenoviruses using SMF concentration has previously been described [7], and were also observed in this study. In addition, SMF failed to detect the low numbers of reads of polyomaviruses and papillomaviruses observed by the other methods.

The larger initial sample volume, and associated organic material and inhibitors, for SMF (10 L) compared with the other methods (4, 1 and 0.2 L for GW, MAF, and PEG, respectively), could be an explanation for the low recovery of the spiked viruses. Inhibitors can affect PCR amplification, quality of the prepared library, and subsequent virus detection. This theory was further supported by the qPCR results were extracts obtained from SMF had a high level of inhibition. Extraction with QIA has previously been shown to impair detection of HAdV in samples with high levels of organic matter [23]. This was also the case in our study, where extraction with QIA inhibited HAdV detection in all cases except when combined with PEG concentration which both had the lowest starting volume (0.2 L) and an additional filtration step.

Sampling volume is an important factor in viral metagenomics, enhancing the sensitivity and increasing the chances of detecting rare viruses. However, in this study, we did not find a positive relation between methods with high sampling volumes and increased sensitivity. This could be due to an increase in inhibitors or other aspects of the employed concentration methods, although this question was not within the scope of this study. Further studies are needed to investigate the influence of sample volumes and viral metagenomics.

In this study, the bioinformatic analyses were done using alignment of single reads to three virus databases, using the program MGmapper. The choice of bioinformatics pipeline can affect results [50] but any biases of our particular approach should be the same on all samples and should therefore not affect the conclusions of this study.

Low levels of MNV were detected in the metagenomics analysis compared to the amounts used for spiking. However, the reasonable high values that could be detected using qPCR, indicated that the initial extraction was successful. Noroviruses have previously been documented to be difficult to detect using metagenomics [51,52] possibly because of the small genome, robust nucleocapsid, or inhibitory RNA secondary structures [53]. Virus species specific extraction efficiency biases are well documented in viral metagenomics [54] and should always be considered when interpreting the results. Quantitative conclusions from viral metagenomics are not possible for all viral species, illustrated by the good correlation between RPM and qPCR data found for HAdV where no correlation was found for MNV.

Several viruses were detected in higher amounts in the negative controls than in the corresponding samples, strongly suggesting them to be procedure contaminants. Contaminating

DNA is a huge challenge for low input metagenomics [24], and contaminating viral nucleotides have previously been detected in polymerases [25], spin columns [27] and DNases [54]. The specific origin of the contaminating viruses in our study was not clear although some avian herpesviruses were only linked to the QIA extracts. The ubiquitous presence of contaminating viruses stress the importance of including negative controls in future viral metagenomics studies, as well as adding measures to reduce the problem [55,56].

When evaluating the efficiency of the tested methods, clear differences were observed. No single method was superior to the others in all of the tested parameters. However, some trends were observed for the concentration methods as PEG scored higher in viral specificity and SMF inhibited detection of both spiked viruses. In the evaluation of the tested extraction methods the NUC kit stood out in regard to viral specificity and RNA virus detection. Nevertheless, if the focus is only on DNA viruses, for example phage studies, NUC might not be the best option since it scored low in viral richness which could result in loss of rare species. Practical aspects of the concentration and extraction methods were not within the scope of this paper, but may also influence the choice of method (S4 and S5 Tables).

In conclusion, we found a significant influence of concentration and extraction protocols on viral richness, viral specificity, viral pathogen detection, and viral community composition for metagenomic analyses of sewage. This is of major importance when interpreting results from the literature and conducting meta-studies. The use of data base resources, such as the European nucleotide archive (ENA) and short read archive (SRA) are also severely hampered by this fact since extraction kit, volume sample, and concentration procedure are not usually included in the metadata of published viromes. We suggest that such metadata will be included in the future, to allow researchers to select and compare studies conducted with similar methodologies.

Supporting Information

S1 Fig. Distribution of reads on kingdom level of the 16 method combinations and their associated negative controls. Samples were processed in triplicate, and the data shown is the average. _S = sample, _C = Negative extraction control. Databases used are listed in S1 Table.
(PDF)

S2 Fig. HAdV concentration measured by qPCR. (A) HAdV concentration in extracts obtained by using four different concentration methods and (B) extraction methods. The bar, box, whiskers and circles represents median, inter-quartile range, inter-quartile range times 1.5, and outliers, respectively. Asterisks represent significance level of a pairwise t-test with "Holm-Bonferroni" adjusted p-values. ** = $p < 0.01$, *** = $p < 0.001$.
(PDF)

S3 Fig. MNV concentrations measured by qPCR. (A) MNV concentration in extracts obtained by using four different concentration methods and (B) extraction methods. The bar, box and whiskers represents the median, the inter-quartile range, and the inter-quartile range times 1.5, respectively. Asterisks represent significance level of a pairwise t-test with "Holm-Bonferroni" adjusted p-values. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.
(PDF)

S4 Fig. Abundance of all detected viral families. Heatmap showing the abundance of all detected viral families, measured in reads per million, in each biological replica for the different method combinations as well as the controls. _S = sample, _C = Negative control.
(PDF)

S1 Table. Sequence information. Number of raw reads, reads after quality assessment, and reads not mapping to PhiX, and thus usable for subsequent analysis. _S = sample, _C = Negative control.
(PDF)

S2 Table. Overview of reference sequence databases and associated download information. Reference sequence information can be obtained from the URL's shown in 'Download information'.
(PDF)

S3 Table. qPCR inhibition of MNV and HAdV. Inhibition of the 16 combinations of concentration and extraction methods. Inhibition was measured using qPCR of undiluted (1:1) and tenfold diluted (1:10) DNA/RNA extracts. The values in the tables represents Δct between the undiluted and 10 fold diluted samples. A Δct = -3.3 represent a perfect 10 fold dilution. Samples marked in red represents undiluted extracts that could not be quantifiable, these samples are regarded as the most inhibited.
(PDF)

S4 Table. Specifications of the four concentration methods applied in this study.
(PDF)

S5 Table. Properties of the four nucleic acid extraction kits applied in this study.
(PDF)

Acknowledgments

We are grateful to Resadije Idrizi for technical assistance and to Dines Thornbjerg and BIOFOS for providing the wastewater.

Author Contributions

Conceptualization: MHH MH XFC.

Formal analysis: MHH OL NT XFC MH.

Funding acquisition: FMA RG CL ACS.

Investigation: MHH MH.

Methodology: MHH MH XFC.

Project administration: MHH MH.

Resources: ACS RG DE MS JFA FMA SBM.

Software: OL NT.

Supervision: MHH MH.

Visualization: MHH MH XFC NT OL.

Writing – original draft: MHH MH XFC NT OL.

Writing – review & editing: MHH MH XFC NT OL MS DE FMA CL SBM JFA RG ACS.

References

1. Newton RJ, Mclellan SL, Dila DK, Vineis JH, Morrison HG, Eren a M, et al. Sewage Reflects the Microbiomes of Human Populations. MBio. 2015; 6: 1–9.
2. Pina S, Buti M, Jardí R, Clemente-Casares P, Jofre J, Girones R. Genetic analysis of hepatitis A virus strains recovered from the environment and from patients with acute hepatitis. J Gen Virol. Microbiology Society; 2001; 82: 2955–63.
3. Hellmer M, Paxeus N, Magnius L, Enache L, Arnholm B, Johansson a., et al. Detection of Pathogenic Viruses in Sewage Provided Early Warnings of Hepatitis A Virus and Norovirus Outbreaks. Appl Environ Microbiol. 2014; 80: 6771–6781. doi: 10.1128/AEM.01981-14 PMID: 25172863
4. Wong K, Fong TT, Bibby K, Molina M. Application of enteric viruses for fecal pollution source tracking in environmental waters. Environ Int. Elsevier B.V.; 2012; 45: 151–164.
5. van den Berg H, Lodder W, van der Poel W, Vennema H, de Roda Husman AM. Genetic diversity of noroviruses in raw and treated sewage water. Res Microbiol. 2005; 156: 532–40. doi: 10.1016/j.resmic.2005.01.008 PMID: 15862452
6. Rosario K, Nilsson C, Lim YW, Ruan Y, Breitbart M. Metagenomic analysis of viruses in reclaimed water. Environ Microbiol. 2009; 11: 2806–2820. doi: 10.1111/j.1462-2920.2009.01964.x PMID: 19555373
7. Cantalupo PG, Calgua B, Zhao G, Hundesa A, Wier AD, Katz JP, et al. Raw sewage harbors diverse viral populations. MBio. 2011; 2.
8. Bibby K, Peccia J. Identification of viral pathogen diversity in sewage sludge by metagenome analysis. Environ Sci Technol. 2013; 47: 1945–51. doi: 10.1021/es305181x PMID: 23346855
9. Ng TFF, Marine R, Wang C, Simmonds P, Kapusinszky B, Bodhidatta L, et al. High variety of known and new RNA and DNA viruses of diverse origins in untreated sewage. J Virol. 2012; 86: 12161–75. doi: 10.1128/JVI.00869-12 PMID: 22933275
10. Bodewes R, van der Giessen J, Haagmans BL, Osterhaus ADME, Smits SL. Identification of multiple novel viruses, including a parvovirus and a hepevirus, in feces of red foxes. J Virol. 2013; 87: 7758–64. doi: 10.1128/JVI.00568-13 PMID: 23616657
11. Daly GM, Bexfield N, Heaney J, Stubbs S, Mayer AP, Palser A, et al. A viral discovery methodology for clinical biopsy samples utilising massively parallel next generation sequencing. PLoS One. 2011; 6.
12. Conceição-neto N, Zeller M, Lefrère H, Bruyn De P, Beller L. Modular approach to customise sample preparation procedures for viral metagenomics : a reproducible protocol for virome analysis. Nat Publ Gr. Nature Publishing Group; 2015; 1–14.
13. John SG, Mendez CB, Deng L, Poulos B, Kauffman AKM, Kern S, et al. A simple and efficient method for concentration of ocean viruses by chemical flocculation. Environ Microbiol Rep. 2011; 3: 195–202. doi: 10.1111/j.1758-2229.2010.00208.x PMID: 21572525
14. Calgua B, Mengewein a., Grunert a., Bofill-Mas S, Clemente-Casares P, Hundesa a., et al. Development and application of a one-step low cost procedure to concentrate viruses from seawater samples. J Virol Methods. 2008; 153: 79–83. doi: 10.1016/j.jviromet.2008.08.003 PMID: 18765255
15. Albinana-Gimenez N, Clemente-Casares P, Calgua B, Huguet JM, Courtois S, Girones R. Comparison of methods for concentrating human adenoviruses, polyomavirus JC and noroviruses in source waters and drinking water using quantitative PCR. J Virol Methods. 2009; 158: 104–109. doi: 10.1016/j.jviromet.2009.02.004 PMID: 19428577
16. Pei L, Rieger M, Lengger S, Ott S, Zawadsky C, Hartmann NM, et al. Combination of Cross fl ow Ultra fi ltration, Monolithic A ffi nity Filtration, and Quantitative Reverse Transcriptase PCR for Rapid Concentration and Quanti fi cation of Model Viruses in Water. 2012;
17. Hurwitz BL, Deng L, Poulos BT, Sullivan MB. Evaluation of methods to concentrate and purify ocean virus communities through comparative, replicated metagenomics. Environ Microbiol. 2013; 15: 1428–40. Available: http://doi.wiley.com/10.1111/j.1462-2920.2012.02836.x doi: 10.1111/j.1462-2920.2012.02836.x PMID: 22845467
18. Kunze A, Pei L, Elsässer D, Niessner R, Seidel M. High performance concentration method for viruses in drinking water. J Virol Methods. Elsevier B.V.; 2015; 222: 132–137.
19. Calgua B, Rodriguez-Manzano J, Hundesa A, Suñen E, Calvo M, Bofill-Mas S, et al. New methods for the concentration of viruses from urban sewage using quantitative PCR. J Virol Methods. 2013; 187: 215–221. doi: 10.1016/j.jviromet.2012.10.012 PMID: 23164995
20. Wesolowska-Andersen A, Bahl MI, Carvalho V, Kristiansen K, Sicheritz-Pontén T, Gupta R, et al. Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. Microbiome. 2014; 2: 19. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4063427&tool=pmcentrez&rendertype=abstract doi: 10.1186/2049-2618-2-19 PMID: 24949196

21. Kennedy NA, Walker AW, Berry SH, Duncan SH, Farquarson FM, Louis P, et al. The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing. PLoS One. 2014; 9: 1–9.
22. Petrich a, Mahony J, Chong S, Broukhanski G, Gharabaghi F, Johnson G, et al. Multicenter comparison of nucleic acid extraction methods for detection of severe acute respiratory syndrome coronavirus RNA in stool specimens. J Clin Microbiol. 2006; 44: 2681–8. doi: 10.1128/JCM.02460-05 PMID: 16891478
23. Iker BC, Bright KR, Pepper IL, Gerba CP, Kitajima M. Evaluation of commercial kits for the extraction and purification of viral nucleic acids from environmental and fecal samples. J Virol Methods. Elsevier B.V.; 2013; 191: 24–30.
24. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 2014; 12: 87. Available: http://www.biomedcentral.com/1741-7007/12/87 doi: 10.1186/s12915-014-0087-z PMID: 25387460
25. Newsome T, Li B-J, Zou N, Lo S-C. Presence of Bacterial Phage-Like DNA Sequences in Commercial Taq DNA Polymerase Reagents. J Clin Microbiol. 2004; 42: 2264–2267. doi: 10.1128/JCM.42.5.2264-2267.2004 PMID: 15131208
26. Naccache SN, Greninger AL, Lee D, Coffey LL, Phan T, Rein-Weston A, et al. The perils of pathogen discovery: origin of a novel parvovirus-like hybrid genome traced to nucleic acid extraction spin columns. J Virol. 2013; 87: 11966–77. doi: 10.1128/JVI.02323-13 PMID: 24027301
27. Rosseel T, Pardon B, De Clercq K, Ozhelvaci O, Van Borm S. False-positive results in metagenomic virus discovery: A strong case for follow-up diagnosis. Transbound Emerg Dis. 2014; 61: 293–299. doi: 10.1111/tbed.12251 PMID: 24912559
28. Wunderlich A, Torggler C, Elsässer D, Lück C, Niessner R, Seidel M. Rapid quantification method for Legionella pneumophila in surface water. Anal Bioanal Chem. Springer Berlin Heidelberg; 2016; 408: 2203–2213.
29. Millen HT, Gonnering JC, Berg RK, Spencer SK, Jokela WE, Pearce JM, et al. Glass Wool Filters for Concentrating Waterborne Viruses and Agricultural Zoonotic Pathogens. J Vis Exp. 2012; 6–11.
30. Leeuwen Van M, Williams MMW, Simon JH, Smits SL, Albert DM, Osterhaus E, et al. Human picobirnaviruses identified by molecular screening of diarrhea samples. J Clin Microbiol. 2010; 48: 1787–94. doi: 10.1128/JCM.02452-09 PMID: 20335418
31. Nordahl Petersen T, Rasmussen S, Hasman H, Carøe C, Bælum J, Charlotte Schultz A, et al. Metagenomic analysis of toilet waste from long distance flights; a step towards global surveillance of infectious diseases and antimicrobial resistance. Sci Rep. Nature Publishing Group; 2015; 5: 11444.
32. Rawsthorne H, Phister TG, Jaykus LA. Development of a fluorescent in situ method for visualization of enteric viruses. Appl Environ Microbiol. 2009; 75: 7822–7827. doi: 10.1128/AEM.01986-09 PMID: 19854924
33. Hernroth BE, Girones R, Allard AK, Hernroth BE, Girones R, Allard AK. Environmental Factors Influencing Human Viral Pathogens and Their Potential Indicator Organisms in the Blue Mussel, Mytilus edulis : the First Scandinavian Report Environmental Factors Influencing Human Viral Pathogens and Their Potential Indicator Organ. Appl Environ Microbiol. 2002; 68: 4523.
34. Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey H a, Ganem D, et al. Microarray-based detection and genotyping of viral pathogens. Proc Natl Acad Sci U S A. 2002; 99: 15687–92. doi: 10.1073/pnas.242579699 PMID: 12429852
35. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011; 17: 10. Available: http://journal.embnet.org/index.php/embnetjournal/article/view/200/479
36. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics. Oxford University Press; 2010; 26: 589–95. Available: http://bioinformatics.oxfordjournals.org/content/26/5/589.full
37. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25: 2078–9. doi: 10.1093/bioinformatics/btp352 PMID: 19505943
38. Bunge J, Woodard L, Böhning D, Foster JA, Connolly S, Allen HK. Estimating population diversity with CatchAll. Bioinformatics. Oxford University Press; 2012; 28: 1045–7.
39. R Core Team. R: A language and environment for statistical computing. R Found Stat Comput Vienna, Austria. Vienna, Austria; 2016; https://www.R-project.org/
40. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. Scand J Stat Scand J Stat. 1979; 6: 65–70. Available: http://www.jstor.org/stable/4615733
41. Vu VQ. ggbiplot: A ggplot2 based biplot [Internet]. 2011. http://github.com/vqv/ggbiplot
42. Kolde R. pheatmap: Pretty Heatmaps [Internet]. 2015. https://cran.r-project.org/package=pheatmap

43. Wagg C, Bender SF, Widmer F, van der Heijden MGA. Soil biodiversity and soil community composition determine ecosystem multifunctionality. Proc Natl Acad Sci U S A. 2014; 111: 5266–70. doi: 10.1073/pnas.1320054111 PMID: 24639507

44. Bibby K, Viau E, Peccia J. Viral metagenome analysis to guide human pathogen monitoring in environmental samples. Lett Appl Microbiol. 2011; 52: 386–92. doi: 10.1111/j.1472-765X.2011.03014.x PMID: 21272046

45. Yang J, Yang F, Ren L, Xiong Z, Wu Z, Dong J, et al. Unbiased parallel detection of viral pathogens in clinical samples by use of a metagenomic approach. J Clin Microbiol. 2011; 49: 3463–9. doi: 10.1128/JCM.00273-11 PMID: 21813714

46. Glass RI, Parashar UD, Estes MK. Norovirus gastroenteritis. N Engl J Med. 2009; 361: 1776–85. doi: 10.1056/NEJMra0804575 PMID: 19864676

47. Tate JE, Burton AH, Boschi-Pinto C, Steele AD, Duque J, Parashar UD, et al. 2008 estimate of worldwide rotavirus-associated mortality in children younger than 5 years before the introduction of universal rotavirus vaccination programmes: a systematic review and meta-analysis. Lancet Infect Dis. Elsevier; 2012; 12: 136–41.

48. Stanaway JD, Flaxman AD, Naghavi M, Fitzmaurice C, Vos T, Abubakar I, et al. The global burden of viral hepatitis from 1990 to 2013: findings from the Global Burden of Disease Study 2013. Lancet. 2016;

49. Scallan E, Hoekstra RM, Angulo FJ, Tauxe R V., Widdowson M-A, Roy SL, et al. Foodborne Illness Acquired in the United States—Major Pathogens. Emerg Infect Dis. 2011; 17: 7–15. doi: 10.3201/eid1701.P11101 PMID: 21192848

50. Borozan I, Watt SN, Ferretti V. Evaluation of Alignment Algorithms for Discovery and Identification of Pathogens Using RNA-Seq. 2013;

51. Bibby K, Peccia J. Identification of Viral Pathogen Diversity in Sewage Sludge by Metagenome Analysis. 2013;

52. Mee ET, Preston MD, Participants S, Minor PD, Schepelmann S, Huang X, et al. Development of a candidate reference material for adventitious virus detection in vaccine and biologicals manufacturing by deep sequencing. Vaccine. 2016; 34: 2035–2043. doi: 10.1016/j.vaccine.2015.12.020 PMID: 26709640

53. Ding Y, Tang Y, Kwok CK, Zhang Y, Bevilacqua PC, Assmann SM. In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. Nature. Nature Publishing Group; 2013; 505: 696–700.

54. Li L, Deng X, Mee ET, Collot-teixeira S, Anderson R, Schepelmann S, et al. Comparing viral metagenomics methods using a highly multiplexed human viral pathogens reagent. J Virol Methods. Elsevier B. V.; 2015; 213: 139–146.

55. Kjartansdóttir KR, Friis-Nielsen J, Asplund M, Mollerup S, Mourier T, Jensen RH, et al. Traces of ATCV-1 associated with laboratory component contamination. Proc Natl Acad Sci. National Academy of Sciences; 2015; 112: E925–E926.

56. Woyke T, Sczyrba A, Lee J, Rinke C, Tighe D, Clingenpeel S, et al. Decontamination of MDA reagents for single cell whole genome amplification. PLoS One. 2011; 6.

ARTICLE III

Metagenòmica per a l'estudi de la contaminació viral en aigua residual urbana

3. Metagenòmica per a l'estudi de la contaminació viral en aigua residual urbana

Amb l'objectiu d'estudiar els virus excretats per la població s'han aplicat tècniques de seqüenciació massiva en mostres d'aigua residual recollides en tres estacions de l'any diferents.

Per a concentrar els virus presents en l'aigua residual es va utilitzar el mètode detallat a Cantalupo i col·laboradors (2011) amb lleugeres modificacions, partint de 10L d'aigua residual urbana concentrada per floculació orgànica amb llet descremada. L'aplicació d'aquest mètode va permetre detectar multitud de virus patògens pertanyents a 37 famílies virals diferents. Els adenovirus humans (HAdV), molt prevalents a l'aigua residual i utilitzats com a indicadors de la contaminació fecal humana, no van poder-se detectar mitjançant les tècniques de seqüenciació massiva, mentre que les mateixes mostres si que van ser quantificades per qPCR específica d'HAdV. Amb l'objectiu de millorar la sensibilitat de les tècniques i d'estudiar de forma específica la família *Adenoviridae* amb tècniques de seqüenciació massiva, es van desenvolupar uns encebadors àmpliament degenerats que amplifiquen la regió de l'hexó d'adenovirus. L'aplicació d'aquest assaig d'amplificació dirigida va demostrar que la majoria de les seqüències corresponien a adenovirus murí i a adenovirus humans de l'espècie F. Per tal de solucionar els problemes de sensibilitat detectats, es va comparar un protocol de concentració basat en ultracentrifugació amb un protocol basat en floculació amb llet descremada però que utilitza un menor volum de mostra inicial. La combinació dels dos protocols va permetre la detecció de 41 famílies víriques, incloent la detecció d'espècies patogèniques pertanyents a les famílies *Caliciviridae, Adenoviridae, Astroviridae*, *Picornaviridae, Polyomaviridae, Papllomaviridae* i *Hepeviridae.* Els dos protocols van demostrar la seva utilitat presentant resultats comparables.

Aquest estudi també inclou una avaluació de la contribució de l'orina en el viroma de l'aigua residual, analitzant els virus presents en una mostra composta d'orina de 14 individus. Els resultats senyalen que l'aportació vírica de l'orina al viroma de l'aigua residual estaria restringida a famílies de virus ADN i en especial, a la família *Polyomaviridae i Papillomaviridae*. Per tant, les principals entrades de virus humans a l'aigua residual es farien a partir de la femta.

L'aigua residual urbana està constituïda per l'excreció de milers d'habitants, constituint una mostra representativa per a desenvolupar estudis epidemiològics. L'estudi del metagenoma víric de l'aigua residual pot donar informació sobre espècies víriques circulant en la població, i per tant, esdevenir informació rellevant de cara a estudis en Salut pública.

Metagenomics for the study of the viral contamination in urban wastewater

Authorship

X.Fernandez-Cassi[1,*,¶,] N. Timoneda[1,2,¶], S. Martínez-Puchol[1], M. Rusiñol[1], J. Rodríguez-Manzano, N. Figuerola[1], R.H. Purcell[3], S. Bofill-Mas[1], J.F. Abril[2], R. Girones[1].

[1] Department of Microbiology, Faculty of Biology, University of Barcelona, Barcelona, Spain.

[2]Computational Genomics Lab, Department of Genetics, Microbiology and Statistics, University of Barcelona; and Institute of Biomedicine (IBUB), University of Barcelona; Barcelona, Catalonia, Spain

[3]*National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA.*

[*] Corresponding author

[¶] These authors contributed equally to this work.

Abstract/summary

The development of a public health viral metagenomics surveillance of the viruses circulating in a population based on the study of urban sewage and contaminating viruses in the environment is of high interest although still in its initial steps. The application of next-generation sequencing (NGS) techniques to study viruses present in urban sewage has been limited to very few studies,, which is in part due to the lack of reliable and sensitive protocols to study viral diversity as well the difficulties with processing NGS data. One important step in the methodology needed is an efficient virus concentration protocol for sewage samples. In this study, different protocols for virus concentration in urban sewage were evaluated. The

application of a concentration method based on organic flocculation of viruses using skimmed milk (SMF) in 10 L of sewage has allowed the detection of many viruses, producing very valuable information on the virome of urban sewage in different seasons, however some viruses as the human Adenovirus (HAdV) were not always detected using the metagenomics approach even when a qPCR assay was positive. In order to evaluate the diversity of adenoviruses present in sewage, a targeted metagenomics assay using one of the previously tested sewage samples by untargeted metagenomics, was studied using general degenerated primers. . The results of the targeted metagenomics assay showed the presence of a high diversity of adenoviral strains, most of them taxonomically assigned to murine Adenoviruses (60%), HAdV-41 (29%) and HAdV-9 (3,6%). In order to increase the sensitivity of the metagenomics assay in urban sewage improvements in the concentration protocols were evaluated. Two different protocols for the virus concentration were comparatively analysed: an ultracentrifugation protocol and a lower-volume SMF protocol (500 ml) producing robust results in the virome with both protocols proving that the bioinformatics pipeline was efficient. The sewage virome presented 41 viral families, including pathogenic viral species that were taxonomically assigned to *Caliciviridae, Adenoviridae, Astroviridae*, *Picornaviridae, Polyomaviridae, Papillomaviridae* and *Hepeviridae.* The contribution of urine to the viral composition of sewage was also evaluated by analysing a pools of urine samples by viral metagenomics and it seems to be restricted to few specific DNA viral families, including Polyomavirus and Papillomavirus species. Amplification of viral strains present in sewage by experimental infections using the Rhesus macaque model allowed for the identification of infective human hepatitis E (HEV) and JC Polyomavirus (JCPyV) but no novel viruses were

identified in the rhesus serum samples. The protocol for the analysis of the virome in urban sewage developed showed to be reliable and the list of species in the sewage virome has been defined including members of the human virome, classical pathogens and emerging strains. A sensitive protocol for the analysis of viruses in sewage and other environmental samples by metagenomics has been proposed. Urban raw sewage consists of the excreta of thousands of inhabitants; therefore, it is a representative sample for epidemiological surveillance purposes. Therefore, the study of the metavirome present in sewage can provide important information of public health significance, highlighting the presence of viral strains circulating within a population while acting as a complex matrix for viral discovery.

Keywords: viral metagenomics, Human adenovirus, viral pathogens, sewage, next-generation sequencing

1. Introduction

In recent years, water scarcity and the application of more sustainable water reuse practices has favoured the use of treated sewage for several purposes, such as crop and green area irrigation, river catchment replenishment and toilet flushing. Conventional treatments applied in wastewater treatment plants are known to be less efficient for viral removal compared to faecal indicator bacteria (FIB) (Gerba et al., 1979; Pusch et al., 2005). This higher viral survival in waste water treatment plants (WWTP) treatments can represent a threat for consumers because WWTP effluents with viruses can contaminate water or food. Raw urban sewage is a complex matrix consisting of urine, faeces and skin desquamation from people. Therefore, raw sewage contains a large variety of viruses, bacteria and protozoa excreted from thousands of inhabitants. Sewage contains pathogenic and commensal viruses, the latter of which might play a beneficial role in the human gut microbiome; also, a high number of plant viruses pass though the human intestines. Sewage additionally contains other non-human inputs, which increases the diversity of this complex ecosystem. Viruses do not have a conserved gene marker, such as 16s, that is shared across all species, hampering the study of viral metagenomes. However, the application of random-primer based sequencing approaches in combination with next-generation sequencing techniques (NGS) has opened a new path for viral discovery, increasing the viral species described each year. The application of viral metagenomics in sewage constitute an excellent tool to monitor and identify potentially known and unknown viral pathogens circulating among the population, contributing to public health surveillance.

Although some viral metagenomics protocols are available for clinical samples (Kohl et al., 2015), only a few manuscripts describe the application

of metagenomics approaches to analyse the viruses present in sewage (Cantalupo et al., 2011; Ng et al., 2012). Previous studies have shown that viruses prevalent in sewage are not always detected by metagenomics, suggesting that protocols should be improved to increase sensitivity. For example, HAdV were hardly detected by Cantalupo and collaborators by NGS, although they had high genome numbers by qPCR.

In the present manuscript, we have studied the diversity of viruses present in raw sewage by testing samples from three different seasons using metagenomics. The application of this methodology allowed for description of the human virome and evaluation of the sensitivity of the technique using HAdVs as a reference virus. Human Adenoviruses (HAdVs) were selected because their double role as pathogens and as specific human viral faecal indicator (Bofill-Mas et al., 2013).

With this purpose, we compared the performance of untargeted metagenomics to an adenovirus-targeted NGS assay and HAdV qPCR values. To increase the number of different viral species identified in sewage, different protocols for virus concentration in urban sewage were evaluated. and an efficient protocol for the analysis of viruses in sewage and other environmental samples by metagenomics has been proposed.

The application of metagenomics in different human body parts has facilitated the study of viral communities in the oral cavity (Ly et al., 2014), gut (Minot et al., 2011), respiratory tract (Willner et al., 2009), skin (Foulongne et al., 2012), blood (Sauvage et al., 2016) and cerebrospinal fluid (Perlejewski et al., 2015). Viral faecal viromes have been studied in healthy (Minot et al., 2011) and unhealthy patients (Linsuwanon et al., 2015) as well as in domestic animals (Mihalov-Kovács et al., 2014); hence, the viral contribution of faeces to raw sewage seems clear. Of note, the

124 viral communities excreted through urine remain poorly studied (Tasha M
125 Santiago-Rodriguez et al., 2015), which is probably because urine was
126 previously considered a sterile environment. To assess the contribution of
127 urine to the virome of raw sewage and to study its viral composition, viruses
128 in pooled urine samples have been also analysed by metagenomics in this
129 study.

130 The infectivity of known and unknown viral species present in raw sewage
131 has been explored by the intravenous inoculation of a sewage sample to
132 Rhesus monkeys as a potential enrichment step prior the application of the
133 metagenomics approach in the rhesus serum samples.

134 Finally, a tailored protocol to analyse sewage and other environmental
135 samples using metagenomics has been proposed. Bioinformatics-specific
136 parameters were adjusted at different levels and new tools were tested to
137 filter out the best set of raw reads, such as those containing the most
138 informative sequences. Those reads were combined into assembled
139 contigs that were later used to detect the known species genomes present
140 in the samples and the relative abundances of the taxonomic groups found
141 in the species mixture. Similarity searches also provided a basic
142 characterization of the pathogenic species present in the samples.

143 2. Materials and methods

144 2.1 Concentration of viral particles from tested samples

145 2.1.1. Concentration of viral particles from raw sewage using skimmed milk
146 organic flocculation (SMF).

147 Three 10-L samples of raw sewage from a UWWTP in Sant Adrià del Besós
148 were collected in Winter, Spring and Summer 2013. Samples were
149 processed after 2 hours of collection. Viral particles were concentrated by

applying the skimmed milk organic flocculation (SMF) method described by Cantalupo et al. (2011). Free DNA from viral concentrates was removed, nucleic acids (NAs) were extracted, and libraries were prepared as explained in section 2.2.

In a second protocol, the reduction of the volume of the sewage sample was also evaluated in order to reduce inhibitory compounds and interfering materials in the viral concentrate.

Briefly, the SMF-adapted protocol used 500 mL of raw sewage that was preconditioned to a pH of 3.5. A volume of 500 µL of a pre-flocculated skim milk solution at pH 3.5 was added to the samples. After 8 h of stirring, flocks were centrifuged at 8000xg for 40 minutes, and the pellet was suspended in 4 mL of phosphate buffer [vol/vol] (0.2 M Na_2HPO_4 and 0.2 M NaH_2PO_4). The viral concentrate was kept at -80°C until further use.

A third protocol based on ultracentrifugation was evaluated in comparison the 500ml SMF protocol. Two samples of 600 mL of raw sewage were collected. Samples were divided into two aliquots: 500 mL for processing according to the SMF protocol adapted from Calgua et al., 2008, and 42 mL for the ultracentrifugation protocol adapted from Pina et al., 1998a. The ultracentrifugation protocol used 42 mL of sewage that was processed as described by Pina et al., 1998a. The obtained SMF and Ultracentrifugation viral concentrates were filtered in 0.45-µm Sterivex filters. Free DNA was removed, NAs were extracted, and libraries were prepared as explained in section 2.2. For both methodologies, the equivalent of 7 mL of a raw sewage sample was analysed in the final constructed libraries. HAdV qPCR was performed on NA extractions as described in section 2.5.

2.1.2 Concentration of viral particles from urine

To explore the viruses excreted by urine and the contribution from urine to the raw sewage virome, 100 mL of urine from 14 healthy volunteers of different ages and origins (7 males and 7 females from 25-63 years old) although most of them living in Barcelona was collected. Individual urine samples were ultracentrifuged for 1 h at 90,000xg at 4°C. The obtained viral pellets were suspended in 300 µL of PBS1X and kept at -80°C until further use. A pooled sample with 1000 µL of each individual urine viral concentrate was obtained. From the pooled sample, 500 µL was DNAse treated, NAs were extracted, and a library was prepared as explained in section 2.2.

2.2 Free DNA removal, nucleic acid extraction, library preparation and sequencing

In all samples, DNAse treatment was performed with the same conditions. Then, 300 µL of raw sewage viral concentrate was treated with 160 U of Turbo DNase (Ambion Cat nº AM1907, Ambion) to remove non-viral free DNA during 1 h at 37°C. DNAse was inactivated using the provided inactivation reagent and centrifugation at 10,000xg for 1.5 minutes. Treated supernatant was collected and kept at 4°C until nucleic viral acid extraction. Then, 280 µL of viral concentrate was extracted using the Qiagen RNA Viral Mini Kit (cat no. 22906, Qiagen, Valencia, CA, USA) without RNA carrier. NAs were eluted using 60 µL of AVE buffer.

For all samples, libraries were prepared following the same protocol. To detect both RNA and DNA viruses, NAs were retrotranscribed using random

nonamer Primer A (5'-GTTTCCCAGTCACGATANNNNNNNNN'-3) as previously described in Wang et al., 2003. Briefly, RNA templates were reverse transcribed using SuperScript III (cat nº 18080093, Life Technologies) and Primer A, which contains a 17-nucleotide specific sequence followed by 9 random nucleotides for random priming. A second cDNA strand was constructed using Sequenase 2.0 (cat nº USBM70775Y200UN, USB/Affymetrix, Cleveland, OH, USA). To address PCR inhibition, 2 library preparations were constructed using 1:2 dilutions of viral NAs. To obtain sufficient DNA for library preparation, a PCR amplification step using Prime rB (5'-GTTTCCCAGTCACGATANNNNNNNNN'-3) and AmpliTaqGold (cat nº 4311806, Life Technologies, Austin, Texax, USA) was performed. After 10 min at 95°C to activate DNA polymerase, the following PCR program was applied: 30 s at 94°C, 30 s at 40°C, 30 s at 50°C for 25 cycles for the ultracentrifugation and low-volume adapted SMF and 40 cycles for the 10L SMF protocol, and finally 60 s at 72°C. PCR products were cleaned and concentrated in a small volume (15 µL) using Zymo DNA clean and concentrator (D4013, Zymo research, USA). Amplified DNA samples were quantified by Qubit 2.0 (cat nº Q32854, Life Technologies, Oregon, USA), and libraries were constructed using a Nextera XT DNA sample preparation kit (Illumina Inc) according to the manufacturer's instructions. Samples were sequenced on Illumina MiSeq 2x250 bp and 2x300 bp, producing paired end reads.

2.3. Bioinformatic pipeline and quality filtering

The quality of raw and clean read sequences was assessed using FASTX-Toolkit software, version 0.0.14 (Hannon Lab). Read sequences were

cleaned using Trimmomatic, version 0.32 (Bolger et al., 2014), taking care of sequencing adaptors and linker contamination. Low quality ends were trimmed considering an average threshold Phred score above Q15 over a running-window of 4 nucleotides. Low complexity sequences, which were mostly biased to repetitive sequences that affect the performance of downstream computational procedures, were then discarded after estimating a linear model based on Trifonov's linguistic complexity (Sarma et al., 1990) and the sequence string compression ratio. Discrimination criteria for the linear model assume low complexity scores below a line with a 45° slope and crossing at 5% below the complexity inflexion point found by the model, which is specific to each sequence set. Finally, duplicated reads were removed in a subsequent step to speed up the downstream assembly. Virome reads were assembled using 90% identify over a minimum of 50% of the read length using CLC Genomics Workbench 4.4 (CLC bio USA, Cambridge, MA), and the resulting contig spectra were used as the primary input for the index. After that, contigs longer than 100 bp were queried for sequence similarity using BLASTN and BLASTX (Altschul et al., 1997, 1990) against the NCBI viral complete genomes database (Brister et al. 2015), the viral division from GenBank nucleotide database (Benson et al., 2015), and viral protein sequences from Uniprot (UniProt Consortium 2015). Species nomenclature and classification was performed according to the NCBI Taxonomy database (Baltimore, 1971) standards. HSPs considered for taxonomical assessment must have an E-value of 10^{-5} and minimum length of 100 bp. Based on the best BLAST result and 90% coverage cut-off, sequences were classified into their likely taxonomic groups of origin. Contigs were merged by Geneious software assembler (Geneious 9, Kearse et al. 2012), and scaffold sequences were subsequently mapped using Geneious mapper tool. Phylogenetic trees were constructed for selected

alignments using Geneious software, and the neighbour-joining method was chosen with 1000 bootstrap replicates. Tables summarizing the number of sequences from the assembly matching each taxonomic unit were built. From those tables, richness ratios were calculated by Catchall software, version 4.0 (Allen et al., 2013); among all the models included in the package, the non-parametric model Chao1 was chosen, which was the model providing the best results on the data-sets. Heatmaps were created using heatmaps from the ggplot2 R graphics library (Kolde, 2015).

2.4 HAdV qPCR as a faecal indicator marker

HAdV were quantified in the urban sewage samples by qPCR as described in previous studies ,(Bofill-Mas et al., 2006).

2.5 Targeted metagenomics for the characterization of adenovirus

To detect and typify all *Mastadenovirus* and other potential adenoviruses present in raw sewage, general primers for AdV hexon were designed. To do so, the hexon region from 149 AdV genomes, recognized by the Adenovirus taxonomy group and retrieved from GenBank, was analysed. The hexon region was selected considering its versatility as a very conserved/variable region (Hernroth et al., 2002), conserved for the design of common primers and variable in the internal sequences useful for typification. Due to the specific requirements of the Roche 454 Junior GS protocol, designed primers were flanked by an adaptor and key sequences to identify samples. Primers and conditions for *Adenovirus* PCR are presented in Supplementary material 3. PCR products were purified using Zymo clean and concentrator (cat nº D4013, Zymo Research). Purified amplicons were then pyrosequenced in a 454 GS Junior System (Life Science-Roche). Obtained raw reads in SFF were transformed to FASTQ using sff_extract from Roche. Adaptors were removed by cutadapt (Martin, 2011a); the complexity and quality of reads were assessed by PrintSeq and

FastQC (Schmieder and Edwards, 2011), which were then trimmed using FASTX-Toolkit software, version 0.0.14 (Hannon Lab). To define non-redundant Operational Taxonomic Units (OTUs), CD-Hit was used and tested at different distance levels from which 0.02 was chosen. A local database was built that contained the hexon region of 153 Adenovirus genomes available from GenBank (2016) and representing different species within the 5 Adenoviridae genus: *Aviadenovirus* (9), *Atadenovirus* (12), *Mastadenovirus* (122), *Siadenovirus* (4) and *Ichtadenovirus*(1). OTUs that matched the 0.02 criteria were blasted against the Adenovirus local database using BLASTN (Altschul et al., 1997, 1990). A phylogenetic tree using Raxml with 1000 bootstrap replicates was computed using Geneious (Geneious 9, Kearse et al. 2012).

2.4. Virus amplification by experimental infection

In collaboration with Dr. Robert H. Purcell (Hepatitis Viruses Section, Laboratory of Infectious Diseases, NIAID, NIH, USA), experimental infections of two Rhesus macaques (*Macaca mulatta*) that were previously immunized for HAV were carried out as a part of a wider study in Bioqual, Rockville, MD in compliance with the guidelines of Bioqual's and NIAID'S Institutional Animal Care and Use Committees. The rhesus macaques were inoculated intravenously with 27 mL of 0.45-µm filtered raw sewage from Barcelona mixed with 3 mL of 10X PBS. Blood from both Rhesus macaques was extracted on a weekly basis over two months to study the potential replication of human viruses present in raw sewage. A blood sample from both animals was extracted a week before inoculation of raw sewage that was used as a negative control.

Nucleic acids and libraries were processed according to section 2.1.2. In total, the following 4 different library preparations were sequenced: a pooled library prior to raw sewage inoculation from the two rhesus monkeys (PW1), two different libraries from each of the animals one week after inoculation (RW1 and RW2), and a pooled library from both rhesus monkeys 4 weeks after inoculation (RW4). Free DNA from viral concentrates was removed, NAs were extracted, and libraries were prepared as described in section 2.2.

3. Results and discussion

3.1. Mi-Seq run outputs in 10L sewage samples from 3 different seasons

Mi-Seq results obtained for sequenced samples are summarized in Supplementary material 1. The virome of urban sewage collected in three different seasons, February, May, September was analysed using 10 L samples of raw sewage, and 37 different viral families were identified. The numbers of different viral species assigned to a given viral family are graphically presented in Figure 1.

Bacteriophage families *Siphoviridae, Myoviridae, Podoviridae* and *Microviridae* show a higher diversity degree in urban sewage. This finding agrees with other studies that bacteriophages are the most abundant organisms on earth (Clokie et al., 2011). The ssDNA Parvoviruses, closely followed by Picornaviruses, are viral families infecting animals/humans with higher diversity. Viral plant *Virgaviridae* species were also abundantly represented in samples. Important human viral pathogens that are taxonomically assigned to *Astroviridae, Caliciviridae*, *Hepeviridae* and *Polyomaviridae* were detected too. Also reads related to viruses belonging to the *Circoviridae* and *Picobirnaviridae* families were sequenced. A summary of the number of reads and contigs associated with those viral

337 families can be found in Table 1. A complete list of detected viral sequences
338 is provided as Supplementary material 2.

339 A wide diversity and abundance of human and animal astroviruses were
340 detected in the winter sample. The majority of the reads from that sample
341 belonged to the MAstV-1 genogroup, whereas MAstV-6, 8 and 9 were less
342 frequent. Similarly, more sequences taxonomically assigned to the
343 *Caliciviridae* viral family, and specifically to different norovirus GI and GII
344 species and human sapoviruses, were detected in winter. The seasonality
345 of Astroviruses and Caliciviruses during low-temperature seasons has been
346 well-documented (Bosch et al., 2014; Haramoto et al., 2006). Within the
347 *Picornaviridae* family, several human and animal Picornaviruses were
348 sequenced, including the recently described Human Salivirus/Klassevirus,
349 several Aichi viruses, and the recently described genus *Cosavirus*. Aichi virus
350 read counts were higher during summer compared to the other tested
351 seasons. Human Enteroviruses from species A, B, C and D had similar
352 numbers, irrespective of the analysed season. Important viral pathogens
353 causing hepatitis transmitted through the consumption of water/food
354 contaminated with faecal material, such as HAV and HEV, were only
355 detected in low numbers in the winter sample although this will be related
356 to the low prevalence of these infections in the studied area..

357 Viral faecal markers present in urban raw sewage, such as human
358 Adenoviruses, were not detected by the metagenomics approach when the
359 10-L SMF protocol was applied. This contrasts with the detection of human
360 Adenoviruses by conventional qPCR in the three seasons tested, Winter,
361 Spring and Summer, with $3{,}18 \times 10^4$ GC/L, $5{,}32 \times 10^5$ GC/L and $1{,}23 \times 10^5$ GC/L,
362 respectively.

363

364 3.1.2. Targeted metagenomics for Adenovirus characterization

365 To address the lack of detection of HAdV and to study the diversity of the
366 genus *Mastadenoviridae* in raw sewage, a target enrichment assay using
367 broadly degenerated primers for the *hexon* region was conducted.
368 Previously concentrated SMF from summer was used because it contained
369 higher numbers of genome copies of HAdV. A total of 55,903 raw reads
370 were generated by pyrosequencing. All raw reads passed the cleaning cut-
371 offs and were used for subsequent analyses. A sequence similarity of 98%
372 was chosen as a cut-off for the homology searches, which resulted in a total
373 of 3,677 different OTUs, accounting for 52,370 sequences from the sample
374 (93.7%). The obtained OTUs were blasted against the custom-built
375 Adenovirus database, falling into 52 phylogenetically different AdV taxons,
376 HAdV A, B, C, D, F and G-. Detected AdVs from raw sewage are shown in
377 Figure 2, and a complete list detailing the abundance of detected AdV is
378 available in Supplementary material 3. Most of the sequences were
379 assigned to Murine Adenovirus-2 (60%) as well as to HAdV from species F,
380 such as to HAdV-41 (29%) and HAdV-40 (0.7%). HAdV-9, from species type
381 D, was the second most abundant HAdV, accounting for a total 3.6% of the
382 reads. The degenerated primers facilitated the detection of a wide range of
383 AdVs with a high variability of the hosts. However, given that some of the
384 detected sequences were from AdV exotic animals and that they clustered
385 with other well-known HAdV species, the used AdV database might not
386 reflect the true diversity within the *Adenoviridae* family, and other excreted
387 human/non-human adenoviruses are yet to be discovered. This is
388 exemplified by some of the detected Simian Adenovirus (SAdV), which is
389 closely related to HAdV40 and 41 (see Figure 2). Therefore, detected SAdV
390 could be variants of the closely related HAdV40 and 41. It should also be
391 considered that in the short region analyzed, few changes are important

392 and errors may be introduced during PCR amplification and sequencing
393 process, it has been described that 454 GS Junior has an overall error rate
394 of 0.18% in a study by Niklas et al(2013),and it is known also that the
395 distribution of errors in the sequences is not homogenous.

396 3.2. Comparative evaluation of ultracentrifugation and SMF of small
397 volumes for virus concentration in sewage
398

399 In order to increase sensitivity, two protocols were comparatively
400 evaluated for the concentration of viruses in sewage and the
401 metegenomics analysis. A modified protocol of virus concentration based
402 on SMF with lower volume of sample, 500ml, and a protocol based on
403 ultracentrifugation. Despite the small volume tested when compared with
404 the results obtained concentrating 10-L of urban sewage, the new modified
405 flocculation protocol allowed the detection of members of viral families
406 that were previously not detected, such as *Adenoviridae, Polyomaviridae*
407 and *Papillomaviridae*, identified when using this methodology.

408 Previous studies in the laboratory did evaluate the effect of different cycle
409 amplifications (25 vs 35 PCR cycles) on the observed viral diversity by
410 estimating the viral richness. Libraries that amplified 35 cycles had a lower
411 average estimated viral richness, affecting the different species of detected
412 dsDNA viruses (data not shown). Both protocols ultracentrifugation and the
413 SMF of 500 ml were used with 25 cycles of amplification before the library
414 construction.

415 Ultracentrifugation is an efficient technique to concentrate viruses, yielding
416 good recoveries. However, the difficulty to simultaneously concentrate viral
417 particles from several samples and the requirement of an ultracentrifuge

device hampers its applicability. A recent comparative study published in collaboration with Hjelmsø et al., (2017) showed that the analysis of 10-L SMF, as described in section 2.1.1, in combination with QIAgen extraction columns (Iker et al., 2013) had inhibition problems, as evidenced by HAdV qPCR quantifications. This observed inhibition might have affected the subsequent detection of some viral species by NGS. A simplified version of SMF using an initial smaller volume that avoided the ultracentrifugation step was compared against the reference ultracentrifugation protocol developed by Pina et al., (1998) to improve and minimize the observed limitations of the reference protocol. Both protocols assayed the same two collected sewage samples, testing the same volume of 7 raw sewage millilitre equivalents per library.

In this comparative study of two samples tested each one in parallel with both methods, producing very good results with the detection of a wide variety of RNA and DNA pathogens with a light increase in the number of sequences (bacteriophages principally) when using the ultracentrifugation method. The four viral concentrates were analysed for HAdV by qPCR showing $8.14x10^5$ GC/L, $1.23x10^5$ GC/ L, $2.19x10^5$GC/L, and $1.48x10^5$ GC/L for HAdV in SMF1, SMF2, Ultra1, and Ultra2, respectively. MI-seq results are summarized in Supplementary material 4. The estimated viral richness values were 755.8 (779.4-732.2), 541.0 (559.9-522.1), 1,066.4 (1,089.6-1,043.2), and 1,318.5 (1,345.6-1,290.4) for SMF1, SMF2, Ultra1, and Ultra2, respectively. These results demonstrate a higher estimated viral richness when using ultracentrifugation compared to SMF. In total, 41 different viral families were detected considering all samples. A complete list of the detected viral families is highlighted in Figure 3. The modified SMF protocol with reduction of the sample volume, allowed the detection of 36 different viral families compared to the 38 different viral families detected by

ultracentrifugation. A higher diversity in viral phage families was observed by this methodology, which impacts the calculated richness by significantly increasing it (see Figure 3). Higher estimated viral richness was also reflected in the detection of few viral human species showing a low number of contigs only detected by the ultracentrifugation protocol, such as the *Anelloviridae*, *Alloherpesviridae, Geminiviridae, Hepeviridae*, *Totiviridae, Geminiviridae,* and *Polyomaviridae* families. Other viral families, such as the *Luteoviridae, Nanoviridae,* and *Baculoviridae* families, were only detected using SMF. For most important viral families, including human pathogenic viruses, such as *Adenoviridae, Caliciviridae, Parvoviridae, Circoviridae, Astroviridae, and Picornaviridae,* a high diversity of viral species was detected with similar results by both SMF-500ml and ultrafiltration protocols, demonstrating the suitability of these concentration methods for the detection of pathogens such as Caliciviruses, the main virus responsible for gastroenteritis outbreaks (Ahmed et al., 2014). Hence, the availability of an effective concentration method to detect pathogenic viruses is crucial if NGS metagenomic data will be used for surveillance purposes. The efficacy of SMF to concentrate ssRNA+ viral particles agrees with previously published results by Hjelmsø et al.(2017). The low-volume SMF protocol allowed for detection of a previously undetected family, *Adenoviridae.* Human Adenoviruses detected by untargeted metagenomics were taxonomically assigned to Human Adenovirus F species (HAdV40 and HAdV41). Higher sensitivity is observed for a specific group of viruses when using targeted metagenomics. The results obtained in urban sewage using the specific targeted metagenomics assay for adenovirus showed a wide diversity of adenoviruses. Murine Adenovirus 2 was found to be the most abundant *Adenoviridae* representative in this specific sample analysed and HAdV 40 and 41 and low numbers of other adenoviral species were also detected. The specific

474 characteristics of the sample and a possible biased preference for Murine
475 Adenoviruses of the highly degenerated Adenovirus *hexon* primers used in
476 the targeted assay may contribute to explain the high number of sequences
477 assigned to this viral species

478 Larger volumes of analysed sample (10 L vs 500 mL) could represent a higher
479 chance to detect rare viral families on sewage. However, larger volumes also
480 have a higher proportion of inhibitors (Schrader et al., 2012). Inhibitors
481 might have affected the PCR amplification step, considering that 40 cycles
482 were needed for 10-L SMF to prepare libraries, while only 25 amplification
483 cycles were needed when 500 mL of SMF was used. Viral metagenomics is
484 limited by the low levels of viral DNA/RNA present in the samples, requiring,
485 in most cases, a PCR-based random amplification step after the RT and
486 sequenase reactions to obtain sufficient DNA for library preparation.
487 Interestingly, viral richness was quite similar despite the different PCR
488 amplification cycles that were applied. The PCR amplification step might
489 introduce bias by amplifying the most abundant genomes such that less
490 abundant genomes might not be sequenced or may be underrepresented
491 (Karlsson et al., 2013). This might be the case for HAdV, as data obtained in
492 previous assays showed that PCR random amplification methods more
493 significantly decreased the estimated viral richness of dsDNA genomes
494 compared to other viral genomes (data not shown). Overall, the data
495 indicate that a concentrated of 500 mL of urban raw sewage is a
496 representative sample volume to study the virome of raw sewage.

497 One of the main objectives of this research was to shed light on the viral
498 families that are present in raw sewage, which we define as the sewage
499 virome. This list should be periodically reviewed using the developed
500 protocols for environmental surveillance and to identify the introduction of

pathogens, novel or emerging viral strains in the population and the environment. A complete list of all different viral species detected in raw sewage using the metagenomics approach in this manuscript is detailed in Supplementary material 2.

In total, more than 11 different viral families considered, or putatively considered, as pathogenic have been detected in raw sewage from Barcelona. *Astroviridae* is a single-stranded, positive-sense RNA viral family of 6.2-7.8 kilobases (kbp) that infects mammals. Human astroviruses (HAstV) are suspected to be involved in 0.5 to 15% of all acute diarrhoea outbreaks in children (Bosch et al., 2014). In the present study, a high diversity of sequences, mainly assigned to the MastV-1 genotype, was detected in all tested samples, but several recombinant genotypes, such as MAstV-6, -8 and -9, were observed in lower abundance. More precisely, the application of NGS techniques has facilitated the detection of these later mentioned animal recombinant astroviruses (Finkbeiner et al., 2009; Kapoor et al., 2009), which are related to neurological disorders in immunocompromised patients (Brown et al., 2015).

Noroviruses (NoV), within the *Caliciviridae* family, are the leading aetiological agent of food-borne disease outbreaks worldwide (Koo et al., 2010). NoV from both genogroups GI and GII were detected in all sewage samples, reflecting a wide diversity within this variable viral family. Those included sequences were taxonomically assigned to NoV GII.4 and NoV GII.17, which are the more frequently reported gastroenteritis genotypes (Chan et al., 2015; Vega et al., 2011). Within the same family, human sapoviruses (HSaV) that belong to GI, GII, GIV, and GV were also found; they have been previously reported as gastroenteritis agents (Oka et al., 2015).

Picornaviridae is a family grouping of more than 30 different genera of

ssRNA+ viruses infecting vertebrates, and it includes important human pathogens, such as hepatitis A virus and poliovirus. Several species of the genera *Kobuvirus*, *Enterovirus (EV)*, *Cosavirus, Salivirus,* and *Cardiovirus* were detected in sewage. Aichi virus (AiV) has been recovered in all seasons and in all tested sewage samples, which agrees with available data (Lodder et al., 2013). Recent studies have suggested that AiV may co-infect with other enteric viruses, causing gastroenteritis (Ambert-Balay et al., 2008; Räsänen et al., 2010). EV is one of the most important genera within the *Picornaviridae* family; it contains 12 different species that infect humans, including *EV* species A to D and *Rhinovirus* species A to C (Plyusnin et al., 2011). Different EV from species A, B, and C and animal enteroviruses from species G and J were also noted. Most of the identified human enteroviruses belong to species A and B, but important enteroviruses from species C, such as Enterovirus-A71, were caught. An increase in enterovirus outbreaks has recently been reported to be caused by emerging recombinant EV strains (Holm-Hansen et al., 2016; Zhang et al., 2010). Other sequences related to the *Salivirus* and *Cosavirus* genera, whose causal role in gastroenteritis is suspected, have been detected (Li et al., 2009; Tseng et al., 2007).

Parvoviridae is a large viral family with a wide range of hosts, from mammals to insects, and constitutes an important component of urban sewage. Several sequences resembling animal parvoviruses that infect dogs, rats, cattle and swine as well as several densoviruses with invertebrate hosts have been identified. Human bocavirus (HBoV) species HBoV1, 2, 3 and, 4 and human bufaviruses have been observed, yet the implications of those parvoviruses in human disease is controversial (Nawaz et al., 2012; Phan et al., 2012), and further studies should be conducted to better characterize their pathogenic role or consider them as part of the human gut viral

556 community.
557 Sequences that are taxonomically assigned to the *Circoviridae* family have
558 been detected in all sewage samples. To date, the Circoviridae family
559 contains two genera, namely, *Circovirus* and *Gyrovirus,* with a third of the
560 proposed genus *Cyclovirus* under revision (Dayaram et al., 2013). Because
561 circoviruses are prevalent in several human fluids, their detection in raw
562 sewage seems reasonable. Their relationship with disease remains unclear,
563 but cycloviruses have been involved in acute nervous system infections (Tan
564 et al., 2013).
565 *Orthohepevirus,* within the *Hepeviridae* family*,* is a genus with the specie
566 Orthoherpesvirus A that includes the viruses causing hepatitis in humans.
567 Genotypes 1 and 2 have been described to infect only humans, while
568 genotypes 3 and 4 are zoonotic (Legrand-Abravanel et al., 2009). The
569 finding in one sample of the HEV genotype 3, frequently detected in swine,
570 illustrates its low prevalence compared to other faecal transmitted viruses
571 causing gastroenteritis (RUTJES et al., 2014).
572 Surprisingly, no members of the Reoviridae family were detected.
573 Important pathogenic viruses within this family include the human
574 Rotaviruses, which are already known as an important gastroenteritis
575 agent in children and cause approximately 453,000 deaths in 2008 (Tate et
576 al., 2012). Although Rotaviruses are detected with similar concentrations
577 as other enteric viruses in sewage (Prado et al., 2011), their prevalence is
578 lower and influenced by seasonality patterns compared to HAdV (El-
579 Senousy et al., 2015; Zhou et al., 2016). Other metagenomic studies failed
580 to detect rotaviruses although they included sewage samples from
581 endemic rotavirus areas (Cantalupo et al., 2011; Ng et al., 2012).
582 *Picobirnaviridae* viruses from the family*,* which also have dsRNA
583 segmented genomes, have been detected in all tested raw sewage

584 samples. Human Picobirnaviruses are prevalent by conventional PCR in
585 100% of sewage samples and have been detected at high concentrations
586 (Symonds et al., 2009). Again, a higher relative abundance of this viral
587 family compared to Rotaviruses should be expected.
588 In the present study, dsDNA viral families, such as *Polyomaviridae*,
589 *Adenoviridae* and *Papillomaviridae*, have been detected. Polyomaviruses
590 and Adenoviruses are excreted by symptomatic and asymptomatic carriers,
591 independent of the seasonality or geographical area. Therefore, they are
592 present in nearly 100% of untreated sewage, which makes them suitable as
593 human viral faecal indicators (Bofill-Mas et al., 2013). Human
594 Papillomaviruses (HPV) have recently been reported in raw sewage (La Rosa
595 et al., 2013). The transmission of papillomaviruses through the
596 consumption of faecal contaminated water or food remains unproven, and
597 further studies on the significance of their molecular detection are needed.
598 Families with insect viruses, such as *Dicistroviridae*, *Iridoviridae,* and
599 *Nodaviridae*, have also been detected, insects could be expected through
600 the sewage system of a city..
601 A high abundance and diversity of plant viruses was found in our samples.
602 Viruses from the *Virgaviridae, Closteroviridae, Partitiviridae,*
603 *Alphaflexiviridae, Betaflexiviridae, Tombusviridae, Bromoviridae,*
604 *Secoviridae, Potyviridae, and Tymoviridae* families seem to be abundant
605 and important components of the sewage virome. Especially diverse are
606 the members of *Virgaviridae* family, which were the second most diverse
607 detected family, irrespective of the concentration method or volume. Plant
608 viruses are highly abundant in human faeces (Zhang et al., 2006); for
609 example, PMMV has been recently related to specific immune responses,
610 fever, and abdominal pains in humans by Colson et al. (Colson et al., 2010).
611 The infectivity of human excreted plant viruses has already been

612 demonstrated (Tomlinson et al., 1982; Zhang et al., 2006). As a result, their
613 presence in WWTP effluents could represent an economic threat for
614 farmers if reclaimed water without a suitable quality control is used for crop
615 irrigation.

616 Bacteriophages were the major fraction from the sewage virome with
617 sequences spotted from *Microviridae*, *Podoviridae*, *Myoviridae*, *Leviviridae*,
618 *Siphoviridae* and *Myoviridae* families. *Microviridae* is the family with a
619 higher level of diversity. Detected phage viral families in the present study
620 agree with other untargeted metagenomic analyses (Tamaki et al., 2012).
621 It is likely that the number of bacteriophages sequences has been
622 underestimated due to the taxonomical assignment of prophages as
623 bacterial DNA.

624 The application of NGS techniques to environmental and clinical samples
625 facilitates the simultaneous analysis of millions of sequences. Of note, a
626 significant fraction of sequences remains unassigned to known taxonomic
627 units after bioinformatics analyses. In the present study, samples were
628 virion-enriched by the applied concentration methods, and the viral
629 concentrate was filtered to remove bacteria, while DNAse was used to
630 remove free DNA. Nevertheless, the percentage of sequences assigned to
631 a known virus taxon was extremely low, but it agreed with previous
632 publications.

633 The evaluated sewage virome is only an initial attempt to address complex
634 water matrices. The lack of a universal viral marker-compared to bacterial
635 16S and the need to sequence all available RNA/DNA present in samples
636 requires concentration methods for viral particles while removing other
637 DNA sources to increase the sensitivity of viral metagenomics. It is expected
638 that the development and availability of improved sequencing

technologies, such as single-molecule nanopore sequencers, in the forthcoming years will provide a more accurate and detailed composition description of the viral mixtures from different types of samples, including those of the sewage virome.

The annotation of the urban sewage virome by applying NGS methods describes the catalogue of the viral species circulating across a given population, which increasingly achieves an important role in public health surveillance. Viruses are more resistant than bacteria to specific treatments applied in WWTP; therefore, they can be present in reclaimed water produced for crop irrigation, surpassing FIB microbiological quality parameters. A previous study by Rosario et al., (2009) demonstrated that reclaimed water contains 1000-fold more virus-like particles than potable water. Although no pathogenic viruses were detected in that study, pathogenic infectious viruses have been detected in reclaimed water in other studies (Rodriguez-Manzano et al., 2012).

3.3. The contribution of urine to the viral composition of sewage

Detected viral sequences from the human urine samples analysed are summarized in Supplementary material 5. The urine viral concentrate contained different DNA viral families infecting humans: *Papillomaviridae, Polyomaviridae,* and sequences distantly related to circular ssDNA families *Circoviridae,* and *Anelloviridae*. Those results highlight that urine contributes to the highly diverse viral composition of urban sewage by introducing principally DNA viruses. Human polyomaviruses were the most abundant, specifcally JC polyomavirus (JCPyV) known to be excreted through urine principally, and with a lower number of sequences BK Polyomavirus (BKPyV),, the 0.76% of the total reads were associated with

666 this family. This excretion route for Polyomaviruses has already been
667 documented in the literature (Egli et al., 2009; Shinohara et al., 1993); for
668 this reason, the group has been widely used as a specific indicator of human
669 excreta in water (Harwood et al., 2009). In recent years, new
670 Polyomaviruses have been described, including up to 13 human
671 Polyomaviruses (Mishra et al., 2014). MCPyV is not excreted through urine
672 (Loyo et al., 2010); instead, it is frequently detected in skin samples in
673 conjunction with human Polyomaviruses 6, 7, and 9 (Foulongne et al.,
674 2012). The lack of detection of the new polyomavirus from urine samples
675 suggests that the excretion patterns of these polyomaviruses might occur
676 through faeces and skin desquamation. Reads of HPV (0.03% of total reads),
677 matching HPV129 and HPV170, which probably come from epithelium
678 desquamation during urination, were identified. The detection of HPVs has
679 been reported in faeces (Di Bonito et al., 2015), raw sewage (La Rosa et al.,
680 2013), and urine (Tasha M. Santiago-Rodriguez et al., 2015). In a prior
681 study, several β-HPV (HPV49, HPV92, and HPV96) and γHPV (HPV121 and
682 HPV178) samples were detected. HPV species detected in this study have
683 not been reported in any of the urine metagenomic studies available to
684 date (Tasha M. Santiago-Rodriguez et al., 2015; Smelov et al., 2016, 2014).
685 Although skin desquamation and excretion through faeces might be the
686 main modes through which human Papillomaviruses land in sewage, the
687 excretion of specific papillomaviruses like the skin-specific γHPV, which
688 might have tropism for the urinary tract, is an interesting finding. Because
689 none of the volunteers participating in this study had been diagnosed with
690 HPV infections or genital warts, HPV could be part of the virome of the
691 urinary tract without causing any known disease. More urine-focused
692 studies, such as those applying specific PCR target enrichment to
693 sequencing, would improve our knowledge of the diversity of HPVs in urine.

694 Sequences distantly related at the protein level to *Circoviridae* and
695 *Anelloviridae* were also observed. ssDNA viruses seem to be ubiquitously
696 present in blood (Vasilyev et al., 2009); therefore, the detection of these
697 specific viral families in urine seems very plausible. With the advent of NGS
698 techniques, there has been a significant increase in viruses classified under
699 these two ssDNA viral families (Kim et al., 2011) and other ssDNA circular
700 viral particles still unclassified (Kim et al., 2012).

701 3.4. Identified infective human viruses present in raw sewage amplified by
702 experimental infection
703

704 One week after inoculation, the first Rhesus monkey presented reads
705 matching JCPyV and Hepatitis E virus, supporting the active replication of
706 these two human viruses identified in urban raw sewage when using animal
707 models. The HEV strain found in the rhesus blood sample was annotated as
708 genotype 3. The inoculation of the environmental HEV strain into Rhesus
709 monkeys is an effective method to replicate the virus (Pina et al., 1998a).
710 Sequences classified within this genotype are frequently reported in the
711 geographical area of the study, in Europe (Clemente-Casares et al., 2009),
712 and this genotype is one of the most commonly detected HEV genotypes in
713 Europe and North America (Clemente-Casares et al., 2003). The second
714 Rhesus monkey did not presented JCPyV or HEV sequences in serum at the
715 studied dates (one week and one month after inoculation). The pooled
716 sample from both Rhesus monkeys after 4 weeks post-inoculation did not
717 have any sequences related to Hepatitis E virus or JCPyV, supporting the
718 model of the acute asymptomatic infection.. Pooled serum samples
719 collected one week before the inoculation showed the presence of several
720 viral plants from the *Virgaviridae* family as well as some phages from the

721 *Microviridae* and *Inoviridae* families. Large fractions of genomic plant DNA
722 have been detected in blood (Spisák et al., 2013), suggesting the possibility
723 that viral DNA/RNA could also be circulating through blood and thus be
724 detected by metagenomics. A total of 1,462 reads (0.08%) in sera sample
725 after inoculation were taxonomically assigned to the *Anelloviridae* family,
726 more specifically human Torque teno virus viruses (TTV) 26 and 27. These
727 two viral species were captured in all Rhesus sera samples, supporting the
728 wide distribution and prevalence of those viruses among mammals (de
729 Villiers and Hausen, 2009). The presence of these viruses in blood has also
730 been reported in humans without any associated disease (Biagini et al.,
731 2013).

732 Conclusions

733 Raw sewage harbours a vast number of different viral families that may
734 contaminate the environment since commonly viruses are not completely
735 removed in WWTPs. The methodology developed based in
736 ultracentrifugation and if a ultracentrifuge is not available, using the SMF
737 protocol for 500 ml samples is useful and produce robust results for the
738 description of the virome of urban sewage detecting both DNA and RNA
739 viruses. The information of the virome of urban sewage may constitute an
740 important data base for known and novel and emerging viral strains
741 excreted in the population in a specific time. Among human viral families,
742 important human pathogens have been detected by NGS, including
743 members of the *Parvoviridae, Caliciviridae, Hepeviridae, Adenoviridae,*
744 *Polyomaviridae, Papillomaviridae,* and *Astroviridae* families. The
745 implementation and application of a low-volume SMF protocol minimized
746 the inhibition problems detected when sampling larger volumes, while
747 offering a representative volume that yielded comparable results to the

748 tested ultracentrifugation method. However, the sensitivity for analysing
749 specific viral groups and reduce representation biases on relatively less
750 abundant viral species is increased by using targeted metagenomics assays
751 designed to amplify specific viral species.

752 The amplification of viruses excreted in sewage through experimental
753 infection in Rhesus macaques allowed for detection of infective HEV and
754 JCPyV from urban sewage showing interesting information on the presence
755 of plant virus in the serum of the macaques and small cDNA viruses still
756 unclassified that will merit further studies.

757 The contribution of urine to sewage seems limited to DNA viral families,
758 mainly to Polyomaviruses JCPyV which appear to be highly excreted and
759 with lower quantities BKPyV,.

760 The use of NGS techniques for sewage analysis can pinpoint major
761 pathogens that circulate in the population and environment, constituting
762 an interesting tool for epidemiologic studies and public health surveillance.

763 Acknowledgements
764 The study reported was partially funded by the Programa RecerCaixa 2012
765 (ACUP-00300) and AGL2011-30461-C02-01/ALI from the Spanish Ministry
766 of Science and Innovation. This study was partially funded by grants from
767 the Catalan Government to Consolidated Research Group VirBaP
768 (2014SRG914) and JPI Water project METAWATER (4193-00001B), with the
769 collaboration of the Institut de Recerca de l'Aigua (IdRA) and the Spanish
770 Adenovirus Network (AdenoNet, BIO2015-68990-REDT) from the
771 Ministerio de Economía y Competitividad of Spain (www.mineco.gob.es).
772 During the development of this study, Xavier Fernandez-Cassi was a fellow

773 of the Catalan Government “AGAUR” (FI-DGR); Natalia Timoneda and
774 Sandra Martínez-Puchol were fellows of the Spanish Ministry of Science.

775 Bibliography

776 Ahmed, S.M., Hall, A.J., Robinson, A.E., Verhoef, L., Premkumar, P.,
777 Parashar, U.D., Koopmans, M., Lopman, B.A., 2014. Global prevalence
778 of norovirus in cases of gastroenteritis: a systematic review and meta-
779 analysis. Lancet Infect. Dis. 14, 725–730. doi:10.1016/S1473-
780 3099(14)70767-4

781 Allen, H.K., Bunge, J., Foster, J.A., Bayles, D.O., Stanton, T.B., 2013.
782 Estimation of viral richness from shotgun metagenomes using a
783 frequency count approach. Microbiome 1, 5. doi:10.1186/2049-2618-
784 1-5

785 Ambert-Balay, K., Lorrot, M., Bon, F., Giraudon, H., Kaplon, J., Wolfer, M.,
786 Lebon, P., Gendrel, D., Pothier, P., 2008. Prevalence and genetic
787 diversity of Aichi virus strains in stool samples from community and
788 hospitalized patients. J. Clin. Microbiol. 46, 1252–1258.
789 doi:10.1128/JCM.02140-07

790 Biagini, P., Bédarida, S., Touinssi, M., Galicher, V., de Micco, P., 2013.
791 Human gyrovirus in healthy blood donors, France. Emerg. Infect. Dis.
792 19, 1014–5. doi:10.3201/eid1906.130228

793 Bofill-Mas, S., Albinana-Gimenez, N., Clemente-Casares, P., Hundesa, A.,
794 Rodriguez-Manzano, J., Allard, A., Calvo, M., Girones, R., 2006.
795 Quantification and stability of human adenoviruses and polyomavirus
796 JCPyV in wastewater matrices. Appl. Environ. Microbiol. 72, 7894–6.
797 doi:10.1128/AEM.00965-06

798 Bofill-Mas, S., Rusiñol, M., Fernandez-Cassi, X., Carratalà, A., Hundesa, A.,
799 Girones, R., 2013. Quantification of human and animal viruses to
800 differentiate the origin of the fecal contamination present in
801 environmental samples. Biomed Res. Int. 2013, 192089.
802 doi:10.1155/2013/192089

803 Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer
804 for Illumina sequence data. Bioinformatics 30, 2114–2120.
805 doi:10.1093/bioinformatics/btu170

806 Bosch, A., Pintó, R.M., Guix, S., 2014. Human astroviruses. Clin. Microbiol.
807 Rev. 27, 1048–1074. doi:10.1128/CMR.00013-14

808 Brown, J.R., Morfopoulou, S., Hubb, J., Emmett, W.A., Ip, W., Shah, D.,
809 Brooks, T., Paine, S.M.L., Anderson, G., Virasami, A., Tong, C.Y.W.,
810 Clark, D.A., Plagnol, V., Jacques, T.S., Qasim, W., Hubank, M., Breuer,
811 J., 2015. Astrovirus VA1/HMO-C: an increasingly recognized
812 neurotropic pathogen in immunocompromised patients. Clin. Infect.
813 Dis. 60, 881–8. doi:10.1093/cid/ciu940

814 Calgua, B., Mengewein, A., Grunert, A., Bofill-Mas, S., Clemente-Casares, P.,
815 Hundesa, A., Wyn-Jones, A.P., López-Pila, J.M., Girones, R., 2008.
816 Development and application of a one-step low cost procedure to
817 concentrate viruses from seawater samples. J. Virol. Methods 153,
818 79–83. doi:10.1016/j.jviromet.2008.08.003

819 Cantalupo, P.G., Calgua, B., Zhao, G., 2011. Raw Sewage Harbors Diverse
820 Viral Populations 2, 1–11. doi:10.1128/mBio.00180-11.Editor

821 Chan, M.C.W., Lee, N., Hung, T.-N., Kwok, K., Cheung, K., Tin, E.K.Y., Lai,
822 R.W.M., Nelson, E.A.S., Leung, T.F., Chan, P.K.S., 2015. Rapid
823 emergence and predominance of a broadly recognizing and fast-

824 evolving norovirus GII.17 variant in late 2014. Nat. Commun. 6, 10061.
825 doi:10.1038/ncomms10061

826 Clemente-Casares, P., Pina, S., Buti, M., Jardi, R., MartIn, M., Bofill-Mas, S.,
827 Girones, R., 2003. Hepatitis E virus epidemiology in industrialized
828 countries. Emerg. Infect. Dis. 9, 448–54. doi:10.3201/eid0904.020351

829 Clemente-Casares, P., Rodriguez-Manzano, J., Girones, R., 2009. Hepatitis
830 E virus genotype 3 and sporadically also genotype 1 circulate in the
831 population of Catalonia, Spain. J. Water Health 7, 664.
832 doi:10.2166/wh.2009.120

833 Clokie, M.R., Millard, A.D., Letarov, A. V, Heaphy, S., 2011. Phages in nature.
834 Bacteriophage 1, 31–45. doi:10.4161/bact.1.1.14942

835 Colson, P., Richet, H., Desnues, C., Balique, F., Moal, V., Grob, J.-J., Berbis,
836 P., Lecoq, H., Harlé, J.-R., Berland, Y., Raoult, D., 2010. Pepper Mild
837 Mottle Virus, a Plant Virus Associated with Specific Immune
838 Responses, Fever, Abdominal Pains, and Pruritus in Humans. PLoS
839 One 5, e10041. doi:10.1371/journal.pone.0010041

840 Dayaram, A., Potter, K.A., Moline, A.B., Rosenstein, D.D., Marinov, M.,
841 Thomas, J.E., Breitbart, M., Rosario, K., Argüello-Astorga, G.R.,
842 Varsani, A., 2013. High global diversity of cycloviruses amongst
843 dragonflies. J. Gen. Virol. 94, 1827–1840. doi:10.1099/vir.0.052654-0

844 de Villiers, E.-M., Hausen, H. zur (Eds.), 2009. TT Viruses, Current Topics in
845 Microbiology and Immunology. Springer Berlin Heidelberg, Berlin,
846 Heidelberg. doi:10.1007/978-3-540-70972-5

847 Di Bonito, P., Della Libera, S., Petricca, S., Iaconelli, M., Sanguinetti, M.,
848 Graffeo, R., Accardi, L., La Rosa, G., 2015. A large spectrum of alpha

849 and beta papillomaviruses are detected in human stool samples. J.
850 Gen. Virol. 96, 607–613. doi:10.1099/vir.0.071787-0

851 Egli, A., Infanti, L., Dumoulin, A., Buser, A., Samaridis, J., Stebler, C., Gosert,
852 R., Hirsch, H.H., 2009. Prevalence of polyomavirus BK and JC infection
853 and replication in 400 healthy blood donors. J. Infect. Dis. 199, 837–
854 46.

855 El-Senousy, W.M., Ragab, A.M.E.-S., Handak, E.M.A.E.H., 2015. Prevalence
856 of Rotaviruses Groups A and C in Egyptian Children and Aquatic
857 Environment. Food Environ. Virol. 7, 132–141. doi:10.1007/s12560-
858 015-9184-6

859 Finkbeiner, S.R., Li, Y., Ruone, S., Conrardy, C., Gregoricus, N., Toney, D.,
860 Virgin, H.W., Anderson, L.J., Vinjé, J., Wang, D., Tong, S., 2009.
861 Identification of a novel astrovirus (astrovirus VA1) associated with an
862 outbreak of acute gastroenteritis. J. Virol. 83, 10836–9.
863 doi:10.1128/JVI.00998-09

864 Foulongne, V., Sauvage, V., Hebert, C., Dereure, O., Cheval, J., Gouilh, M.A.,
865 Pariente, K., Segondy, M., Burguière, A., Manuguerra, J.-C., Caro, V.,
866 Eloit, M., 2012. Human Skin Microbiota: High Diversity of DNA Viruses
867 Identified on the Human Skin by High Throughput Sequencing. PLoS
868 One 7, e38499. doi:10.1371/journal.pone.0038499

869 Gerba, C.P., Goyal, S.M., LaBelle, R.L., Cech, I., Bodgan, G.F., 1979. Failure
870 of indicator bacteria to reflect the occurrence of enteroviruses in
871 marine waters. Am. J. Public Health 69, 1116–9.

872 Haramoto, E., Katayama, H., Oguma, K., Yamashita, H., Tajima, A., Nakajima,
873 H., Ohgaki, S., 2006. Seasonal profiles of human noroviruses and
874 indicator bacteria in a wastewater treatment plant in Tokyo, Japan.

875 Water Sci. Technol. 54, 301–8.

876 Harwood, V.J., Brownell, M., Wang, S., Lepo, J., Ellender, R.D., Ajidahun, A.,
877 Hellein, K.N., Kennedy, E., Ye, X., Flood, C., 2009. Validation and field
878 testing of library-independent microbial source tracking methods in
879 the Gulf of Mexico. Water Res. 43, 4812–4819.
880 doi:10.1016/j.watres.2009.06.029

881 Hernroth, B.E., Conden-Hansson, A.C., Rehnstam-Holm, A.S., Girones, R.,
882 Allard, A.K., 2002. Environmental factors influencing human viral
883 pathogens and their potential indicator organisms in the blue mussel,
884 Mytilus edulis: The first Scandinavian report. Appl. Environ. Microbiol.
885 68, 4523–4533. doi:10.1128/AEM.68.9.4523-4533.2002

886 Hjelmsø, M.H., Hellmér, M., Fernandez-Cassi, X., Timoneda, N.,
887 Lukjancenko, O., Seidel, M., Elsässer, D., Aarestrup, F.M., Löfström, C.,
888 Bofill-Mas, S., Abril, J.F., Girones, R., Schultz, A.C., 2017. Evaluation of
889 Methods for the Concentration and Extraction of Viruses from
890 Sewage in the Context of Metagenomic Sequencing. PLoS One 12,
891 e0170199. doi:10.1371/journal.pone.0170199

892 Holm-Hansen, C.C., Midgley, S.E., Fischer, T.K., 2016. Global emergence of
893 enterovirus D68: A systematic review. Lancet Infect. Dis.
894 doi:10.1016/S1473-3099(15)00543-5

895 Iker, B.C., Bright, K.R., Pepper, I.L., Gerba, C.P., Kitajima, M., 2013.
896 Evaluation of commercial kits for the extraction and purification of
897 viral nucleic acids from environmental and fecal samples. J. Virol.
898 Methods 191, 24–30. doi:10.1016/j.jviromet.2013.03.011

899 Kapoor, A., Li, L., Victoria, J., Oderinde, B., Mason, C., Pandey, P., Zaidi, S.Z.,
900 Delwart, E., 2009. Multiple novel astrovirus species in human stool. J.

Gen. Virol. 90, 2965–2972. doi:10.1099/vir.0.014449-0

Karlsson, O.E., Belák, S., Granberg, F., 2013. The effect of preprocessing by sequence-independent, single-primer amplification (SISPA) on metagenomic detection of viruses. Biosecur. Bioterror. 11 Suppl 1, S227-34. doi:10.1089/bsp.2013.0008

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649. doi:10.1093/bioinformatics/bts199

Kim, H.K., Park, S.J., Nguyen, V.G., Song, D.S., Moon, H.J., Kang, B.K., Park, B.K., 2012. Identification of a novel single-stranded, circular DNA virus from bovine stool. J. Gen. Virol. 93, 635–639. doi:10.1099/vir.0.037838-0

Kim, M.-S., Park, E.-J., Roh, S.W., Bae, J.-W., 2011. Diversity and abundance of single-stranded DNA viruses in human feces. Appl. Environ. Microbiol. 77, 8062–70. doi:10.1128/AEM.06331-11

Kohl, C., Brinkmann, A., Dabrowski, P.W., Radonić, A., Nitsche, A., Kurth, A., 2015. Protocol for metagenomic virus detection in clinical specimens. Emerg. Infect. Dis. 21, 48–57. doi:10.3201/eid2101.140766

Kolde, R., 2015. pheatmap: Pretty Heatmaps.

Koo, H.L., Ajami, N., Atmar, R.L., DuPont, H.L., 2010. Noroviruses: The leading cause of gastroenteritis worldwide. Discov. Med. 10, 61–70. doi:10.1007/s12560-010-9038-1.Noroviruses

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, e52391. doi:10.1371/journal.pone.0052391

Legrand-Abravanel, F., Mansuy, J.-M., Dubois, M., Kamar, N., Peron, J.-M., Rostaing, L., Izopet, J., 2009. Hepatitis E virus genotype 3 diversity, France. Emerg. Infect. Dis. 15, 110–4. doi:10.3201/eid1501.080296

Li, L., Victoria, J., Kapoor, A., Blinkova, O., Wang, C., Babrzadeh, F., Mason, C.J., Pandey, P., Triki, H., Bahri, O., Oderinde, B.S., Baba, M.M., Bukbuk, D.N., Besser, J.M., Bartkus, J.M., Delwart, E.L., 2009. A novel picornavirus associated with gastroenteritis. J. Virol. 83, 12002–6. doi:10.1128/JVI.01241-09

Linsuwanon, P., Poovorawan, Y., Li, L., Deng, X., Vongpunsawad, S., Delwart, E., 2015. The fecal virome of children with hand, foot, and mouth disease that tested PCR negative for pathogenic enteroviruses. PLoS One 10. doi:10.1371/journal.pone.0135573

Lodder, W.J., Rutjes, S.A., Takumi, K., de Roda Husman, A.M., 2013. Aichi virus in sewage and surface water, the Netherlands. Emerg. Infect. Dis. 19, 1222–1230. doi:10.3201/eid1908.130312

Loyo, M., Guerrero-Preston, R., Brait, M., Hoque, M.O., Chuang, A., Kim, M.S., Sharma, R., Liégeois, N.J., Koch, W.M., Califano, J.A., Westra, W.H., Sidransky, D., 2010. Quantitative detection of Merkel cell virus in human tissues and possible mode of transmission. Int. J. Cancer 126, NA-NA. doi:10.1002/ijc.24737

Ly, M., Abeles, S.R., Boehm, T.K., Robles-Sikisaka, R., Naidu, M., Santiago-Rodriguez, T., Pride, D.T., 2014. Altered oral viral ecology in

952 association with periodontal disease. MBio 5, e01133-14.
953 doi:10.1128/mBio.01133-14

954 Mihalov-Kovács, E., Fehér, E., Martella, V., Bányai, K., Farkas, S.L., 2014. The
955 fecal virome of domesticated animals. Virusdisease 25, 150–7.
956 doi:10.1007/s13337-014-0192-1

957 Minot, S., Sinha, R., Chen, J., Li, H., Keilbaugh, S.A., Wu, G.D., Lewis, J.D.,
958 Bushman, F.D., 2011. The human gut virome: inter-individual
959 variation and dynamic response to diet. Genome Res. 21, 1616–25.
960 doi:10.1101/gr.122705.111

961 Mishra, N., Pereira, M., Rhodes, R.H., An, P., Pipas, J.M., Jain, K., Kapoor, A.,
962 Briese, T., Faust, P.L., Lipkin, W.I., 2014. Identification of a Novel
963 Polyomavirus in a Pancreatic Transplant Recipient With Retinal
964 Blindness and Vasculitic Myopathy. J. Infect. Dis. 210, 1595–1599.
965 doi:10.1093/infdis/jiu250

966 Nawaz, S., Allen, D.J., Aladin, F., Gallimore, C., Iturriza-Gómara, M., 2012.
967 Human bocaviruses are not significantly associated with
968 gastroenteritis: results of retesting archive DNA from a case control
969 study in the UK. PLoS One 7, e41346.
970 doi:10.1371/journal.pone.0041346

971 Nicholas, K. B. and H. B. Nicholas (2006). GeneDoc: Multiple sequence
972 alignment editor & shading utility.

973 Ng, T.F.F., Marine, R., Wang, C., Simmonds, P., Kapusinszky, B., Bodhidatta,
974 L., Oderinde, B.S., Wommack, K.E., Delwart, E., 2012. High Variety of
975 Known and New RNA and DNA Viruses of Diverse Origins in Untreated
976 Sewage. J. Virol. 86, 12161–12175. doi:10.1128/JVI.00869-12

977 Niklas, N., Pröll, J., Danzer, M., Stabentheiner, S., Hofer, K., Gabriel, C.,
978 2013. Routine performance and errors of 454 HLA exon sequencing in
979 diagnostics. BMC Bioinformatics 14, 176. doi:10.1186/1471-2105-14-
980 176

981 Oka, T., Wang, Q., Katayama, K., Saif, L.J., 2015. Comprehensive review of
982 human sapoviruses. Clin. Microbiol. Rev. 28, 32–53.
983 doi:10.1128/CMR.00011-14

984 Perlejewski, K., Popiel, M., Laskus, T., Nakamura, S., Motooka, D., Stokowy,
985 T., Lipowski, D., Pollak, A., Lechowicz, U., Caraballo Cortés, K., Stępień,
986 A., Radkowski, M., Bukowska-Ośko, I., 2015. Next-generation
987 sequencing (NGS) in the identification of encephalitis-causing viruses:
988 Unexpected detection of human herpesvirus 1 while searching for
989 RNA pathogens. J. Virol. Methods 226, 1–6.
990 doi:10.1016/j.jviromet.2015.09.010

991 Phan, T.G., Vo, N.P., Bonkoungou, I.J.O., Kapoor, A., Barro, N., O'Ryan, M.,
992 Kapusinszky, B., Wang, C., Delwart, E., 2012. Acute Diarrhea in West
993 African Children: Diverse Enteric Viruses and a Novel Parvovirus
994 Genus. J. Virol. 86, 11024–11030. doi:10.1128/JVI.01427-12

995 Pina, S., Jofre, J., Emerson, S.U., Purcell, R.H., Girones, R., 1998a.
996 Characterization of a strain of infectious hepatitis E virus isolated from
997 sewage in an area where hepatitis E is not endemic. Appl. Environ.
998 Microbiol. 64, 4485–8.

999 Pina, S., Puig, M., Lucena, F., Jofre, J., Girones, R., 1998b. Viral pollution in
1000 the environment and in shellfish: human adenovirus detection by PCR
1001 as an index of human viruses. Appl. Environ. Microbiol. 64, 3376–82.

1002 Plyusnin, A., Beaty, B.J., Elliott, R.M., Goldbach, R., Kormelink, R., Lundkvist,

1003 A., Schmaljohn, C.S., Tesh, R.B., 2011. Virus Taxonomy: Ninth Report
1004 of the International Committee on Taxonomy of Viruses, vol. 9, in:
1005 Edited by King, AMQ, Adams, J. & Lefkowitz, E.. Amsterdam: Elsevier.
1006 pp. 735–741.

1007 Prado, T., Silva, D.M., Guilayn, W.C., Rose, T.L., Gaspar, A.M.C.,
1008 Miagostovich, M.P., 2011. Quantification and molecular
1009 characterization of enteric viruses detected in effluents from two
1010 hospital wastewater treatment plants. Water Res. 45, 1287–1297.
1011 doi:10.1016/j.watres.2010.10.012

1012 Pusch, D., Oh, D.-Y., Wolf, S., Dumke, R., Schröter-Bobsin, U., Höhne, M.,
1013 Röske, I., Schreier, E., 2005. Detection of enteric viruses and bacterial
1014 indicators in German environmental waters. Arch. Virol. 150, 929–47.
1015 doi:10.1007/s00705-004-0467-8

1016 Räsänen, S., Lappalainen, S., Kaikkonen, S., Hämäläinen, M., Salminen, M.,
1017 Vesikari, T., 2010. Mixed viral infections causing acute gastroenteritis
1018 in children in a waterborne outbreak. Epidemiol. Infect. 138, 1227–
1019 1234. doi:10.1017/S0950268809991671

1020 Rodriguez-Manzano, J., Alonso, J.L., Ferrús, M.A., Moreno, Y., Amorós, I.,
1021 Calgua, B., Hundesa, A., Guerrero-Latorre, L., Carratala, A., Rusiñol,
1022 M., Girones, R., 2012. Standard and new faecal indicators and
1023 pathogens in sewage treatment plants, microbiological parameters
1024 for improving the control of reclaimed water. Water Sci. Technol. 66,
1025 2517–23. doi:10.2166/wst.2012.233

1026 Rosario, K., Nilsson, C., Lim, Y.W., Ruan, Y., Breitbart, M., 2009.
1027 Metagenomic analysis of viruses in reclaimed water. Environ.
1028 Microbiol. 11, 2806–2820. doi:10.1111/j.1462-2920.2009.01964.x

1029 RUTJES, S.A., BOUWKNEGT, M., van der GIESSEN, J.W., de RODA HUSMAN,
1030 A.M., REUSKEN, C.B.E.M., 2014. Seroprevalence of Hepatitis E Virus in
1031 Pigs from Different Farming Systems in The Netherlands. J. Food Prot.
1032 77, 640–642. doi:10.4315/0362-028X.JFP-13-302

1033 Santiago-Rodriguez, T.M., Ly, M., Bonilla, N., Pride, D.T., 2015. The human
1034 urine virome in association with urinary tract infections. Front.
1035 Microbiol. 6, 14. doi:10.3389/fmicb.2015.00014

1036 Santiago-Rodriguez, T.M., Ly, M., Bonilla, N., Pride, D.T., 2015. The human
1037 urine virome in association with urinary tract infections. Front.
1038 Microbiol. 6, 14. doi:10.3389/fmicb.2015.00014

1039 Sarma, R.H., Saram, M.H. (Mukti H.., State University of New York at
1040 Albany., 1990. Structure & methods : proceedings of the Sixth
1041 Conversation in the Discipline Biomolecular Stereodynamics held at
1042 the State University of New York at Albany, June 6-10, 1989. Adenine
1043 Press.

1044 Sato, M., Kuroda, M., Kasai, M., Matsui, H., Fukuyama, T., Katano, H.,
1045 Tanaka-Taya, K., 2016. Acute encephalopathy in an
1046 immunocompromised boy with astrovirus-MLB1 infection detected
1047 by next generation sequencing. J. Clin. Virol. 78, 66–70.
1048 doi:10.1016/j.jcv.2016.03.010

1049 Sauvage, V., Laperche, S., Cheval, J., Muth, E., Dubois, M., Boizeau, L.,
1050 Hébert, C., Lionnet, F., Lefrère, J.-J., Eloit, M., 2016. Viral
1051 metagenomics applied to blood donors and recipients at high risk for
1052 blood-borne infections. Blood Transfus. 14, 400–7.
1053 doi:10.2450/2016.0160-15

1054 Schrader, C., Schielke, A., Ellerbroek, L., Johne, R., 2012. PCR inhibitors -

occurrence, properties and removal. J. Appl. Microbiol. 113, 1014–1026. doi:10.1111/j.1365-2672.2012.05384.x

Shinohara, T., Matsuda, M., Cheng, S.H., Marshall, J., Fujita, M., Nagashima, K., 1993. BK virus infection of the human urinary tract. J. Med. Virol. 41, 301–5.

Smelov, V., Arroyo Mühr, L.S., Bzhalava, D., Brown, L.J., Komyakov, B., Dillner, J., 2014. Metagenomic sequencing of expressed prostate secretions. J. Med. Virol. 86, 2042–2048. doi:10.1002/jmv.23900

Smelov, V., Bzhalava, D., Arroyo Muhr, L.S., Eklund, C., Komyakov, B., Gorelov, a, Dillner, J., Hultin, E., 2016. Detection of DNA viruses in prostate cancer. Sci Rep 6, 25235. doi:10.1038/srep25235

Spisák, S., Solymosi, N., Ittzés, P., Bodor, A., Kondor, D., Vattay, G., Barták, B.K., Sipos, F., Galamb, O., Tulassay, Z., Szállási, Z., Rasmussen, S., Sicheritz-Ponten, T., Brunak, S., Molnár, B., Csabai, I., 2013. Complete Genes May Pass from Food to Human Blood. PLoS One 8, e69805. doi:10.1371/journal.pone.0069805

Symonds, E.M., Griffin, D.W., Breitbart, M., 2009. Eukaryotic viruses in wastewater samples from the United States. Appl. Environ. Microbiol. 75, 1402–1409. doi:10.1128/AEM.01899-08

Tamaki, H., Zhang, R., Angly, F.E., Nakamura, S., Hong, P.-Y., Yasunaga, T., Kamagata, Y., Liu, W.-T., 2012. Metagenomic analysis of DNA viruses in a wastewater treatment plant in tropical climate. Environ. Microbiol. 14, 441–452. doi:10.1111/j.1462-2920.2011.02630.x

Tan, L. Van, van Doorn, H.R., Nghia, H.D.T., Chau, T.T.H., Tu, L.T.P., de Vries, M., Canuti, M., Deijs, M., Jebbink, M.F., Baker, S., Bryant, J.E., Tham,

1080 N.T., BKrong, N.T.T.C., Boni, M.F., Loi, T.Q., Phuong, L.T., Verhoeven,
1081 J.T.P., Crusat, M., Jeeninga, R.E., Schultsz, C., Chau, N.V.V., Hien, T.T.,
1082 van der Hoek, L., Farrar, J., de Jong, M.D., 2013. Identification of a new
1083 cyclovirus in cerebrospinal fluid of patients with acute central nervous
1084 system infections. MBio 4. doi:10.1128/mBio.00231-13

1085 Tate, J.E., Burton, A.H., Boschi-Pinto, C., Steele, A.D., Duque, J., Parashar,
1086 U.D., WHO-coordinated Global Rotavirus Surveillance Network, 2012.
1087 2008 estimate of worldwide rotavirus-associated mortality in children
1088 younger than 5 years before the introduction of universal rotavirus
1089 vaccination programmes: a systematic review and meta-analysis.
1090 Lancet. Infect. Dis. 12, 136–41. doi:10.1016/S1473-3099(11)70253-5

1091 Tomlinson, J.A., Faithfull, E., Flewett, T.H., Beards, G., 1982. Isolation of
1092 infective tomato bushy stunt virus after passage through the human
1093 alimentary tract. Nature 300, 637–8.

1094 Tseng, C.-H., Tsai, H.-J., Kirkwood, C.D., Wang, D., Hay, C., Hallenbeck, P.,
1095 Knowles, N., Lemon, S., Minor, P., Pallansch, M., Palmenberg, A.,
1096 Skern, T., 2007. Sequence analysis of a duck picornavirus isolate
1097 indicates that it together with porcine enterovirus type 8 and simian
1098 picornavirus type 2 should be assigned to a new picornavirus genus.
1099 Virus Res. 129, 104–114. doi:10.1016/j.virusres.2007.06.023

1100 Vasilyev, E. V, Trofimov, D.Y., Tonevitsky, A.G., Ilinsky, V. V, Korostin, D.O.,
1101 Rebrikov, D. V, 2009. Torque Teno Virus (TTV) distribution in healthy
1102 Russian population. Virol. J. 6, 134. doi:10.1186/1743-422X-6-134

1103 Vega, E., Barclay, L., Gregoricus, N., Williams, K., Lee, D., Vinjé, J., 2011.
1104 Novel surveillance network for norovirus gastroenteritis outbreaks,
1105 United States. Emerg. Infect. Dis. 17, 1389–1395.

doi:10.3201/eid1708.101837

Wang, D., Urisman, A., Liu, Y.T., Springer, M., Ksiazek, T.G., Erdman, D.D., Mardis, E.R., Hickenbotham, M., Magrini, V., Eldred, J., Latreille, J.P., Wilson, R.K., Ganem, D., DeRisi, J.L., 2003. Viral discovery and sequence recovery using DNA microarrays. PLoS Biol. 1. doi:10.1371/journal.pbio.0000002

Willner, D., Furlan, M., Haynes, M., Schmieder, R., Angly, F.E., Silva, J., Tammadoni, S., Nosrat, B., Conrad, D., Rohwer, F., 2009. Metagenomic Analysis of Respiratory Tract DNA Viral Communities in Cystic Fibrosis and Non-Cystic Fibrosis Individuals. PLoS One 4, e7370. doi:10.1371/journal.pone.0007370

Zhang, T., Breitbart, M., Lee, W.H., Run, J.-Q., Wei, C.L., Soh, S.W.L., Hibberd, M.L., Liu, E.T., Rohwer, F., Ruan, Y., 2006. RNA viral community in human feces: prevalence of plant pathogenic viruses. PLoS Biol. 4, e3. doi:10.1371/journal.pbio.0040003

Zhang, Y., Zhu, Z., Yang, W., Ren, J., Tan, X., Wang, Y., Mao, N., Xu, S., Zhu, S., Cui, A., Zhang, Y., Yan, D., Li, Q., Dong, X., Zhang, J., Zhao, Y., Wan, J., Feng, Z., Sun, J., Wang, S., Li, D., Xu, W., 2010. An emerging recombinant human enterovirus 71 responsible for the 2008 outbreak of hand foot and mouth disease in Fuyang city of China. Virol. J. 7, 94. doi:10.1186/1743-422X-7-94

Zhou, N., Lv, D., Wang, S., Lin, X., Bi, Z., Wang, H., Wang, P., Zhang, H., Tao, Z., Hou, P., Song, Y., Xu, A., 2016. Continuous detection and genetic diversity of human rotavirus A in sewage in eastern China, 2013–2014. Virol. J. 13, 153. doi:10.1186/s12985-016-0609-0

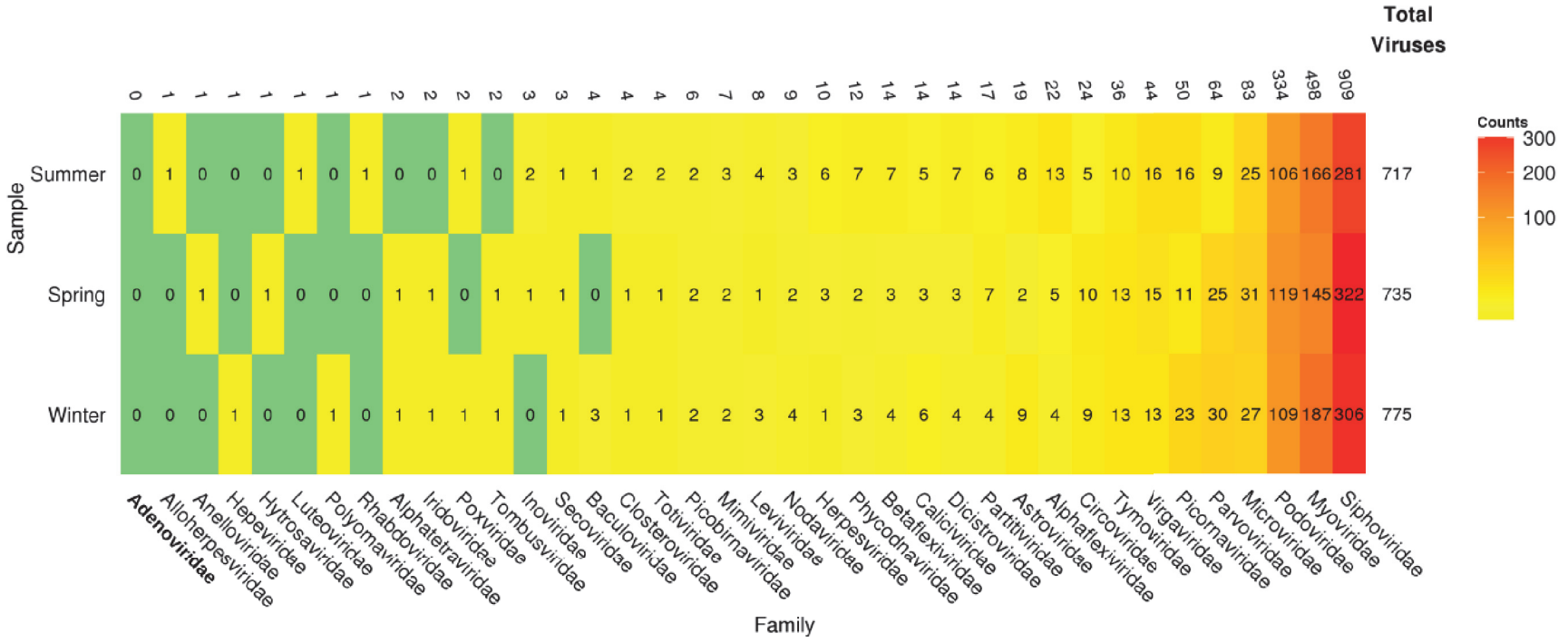
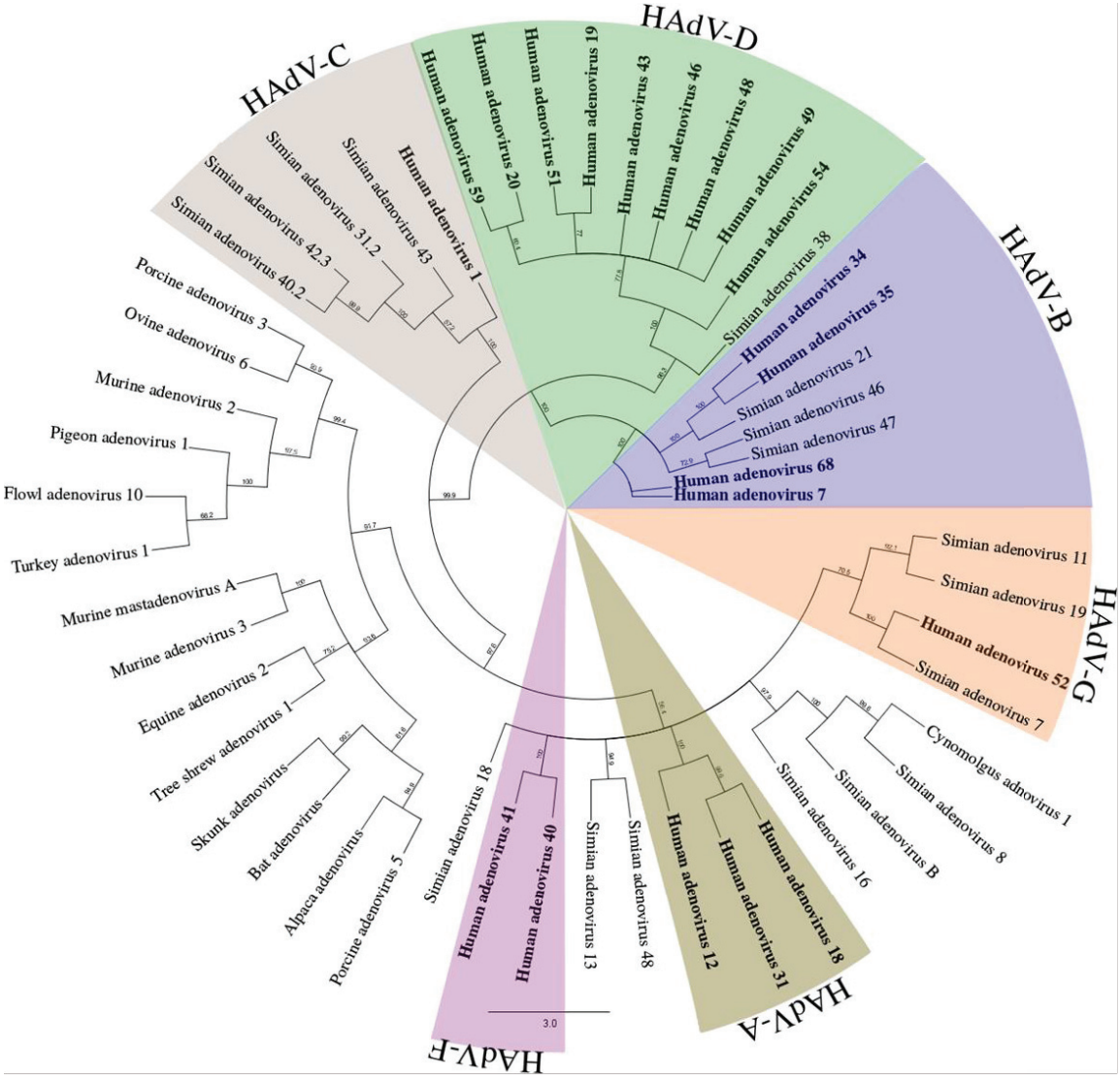


Figure 1. Relative abundance of viral species classified by family. The heatmap shows the relative abundance of 28 viral families detected over 3 different seasonal samples. Numbers within each cell represent the number of sequences that had at least a positive BLAST hit to known species and passed all the selection criteria. The colours range from green (not detected), to red (high relative abundance). Top row and right column correspond to the count sums by category and sample respectively.

Figure 2. Phylogenetic tree based on the complete nucleotide Adenovirus hexon region. Detected Adenovirus sequences from 454 sequencing experiment that presented a match in sewage were aligned. Human adenovirus species are shown in boldface, where colored arcs highlight the distinct taxon groups ranging from HAdV-A to HAdV-F. The tree was built using the neighbor-joining method and 1000 bootstrap replicates (bootstrap values are shown on the tree branches).



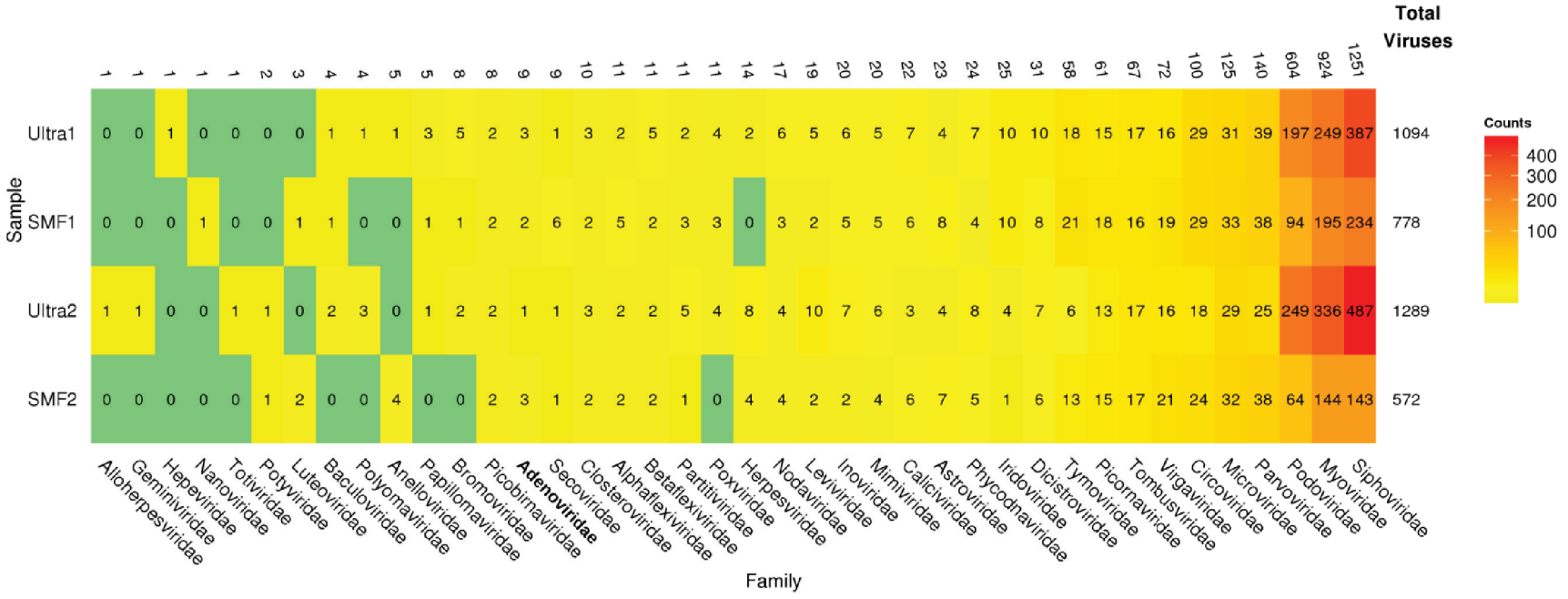


Figure 3. Relative abundance of viral species classified by family. The heatmap shows the relative abundance of 35 viral families detected over 4 different concentration method samples. Numbers within each cell represent the number of sequences that had at least a positive BLAST hit to known species and passed all the selection criteria. The colours range from green (not detected), to red (high relative abundance). Top row and right column correspond to the count sums by category and sample respectively.

Table 1.- Important and potentially pathogenic Human viral families detected in raw sewage by using 10L SMF

Family	Genus/Species	Winter	Spring	Summer	Total hits
Astroviridae	MAstV-1	48	7	5	60
	MAstV-6	8	1	10	19
	MAstV-8/9	18	0	2	20
Parvoviridae	HBoV-2	6	1	0	8
	HBoV-3	3	0	0	3
	HBoV-4	1	2	0	3
	AAV2	12	4	1	17
	AAV3	2	0	0	2
	AAV5	2	0	0	2
Caliciviridae	HSaVGI	7	2	4	13
	HSaVGII	11	0	3	14
	NoVGI	17	6	3	26
	NoVGII	20	1	1	22
Polyomaviridae	JCPyV	4	0	0	4
Circoviridae	Human circovirus	5	12	2	19
Picornaviridae	HAV	2	0	0	2
	Salivirus	20	20	22	62
	Cosavirus	0	2	1	3
	Rosavirus	1	0	0	1
	Enterovirus A	4	1	1	6
	Enterovirus B	10	4	5	19
	Enterovirus C	5	2	1	8
	Enterovirus D	1	0	0	1
	Enterovirus J	0	0	1	1
	Rabovirus	10	0	1	11
	Cardiovirus	8	7	5	20
	Aichi virus	16	12	42	70
Hepeviridae	HEV	4	0	0	4
Picobirnaviridae	Human picobirnavirus	37	11	15	63

Samples	WINTER			SPRING			SUMMER		
	Sequences		Nucleotides	Sequences		Nucleotides	Sequences		Nucleotides
MiSeq Raw READS	2,862,464		653,150,240	1,225,920		290,054,808	951,664		223,053,943
Clean READS									
Pair-Ends	2,820,868		644,330,961	1,172,444		266,982,894	779,690		172,598,345
Single-Ends	7,351		1,802,212	9,158		1,676,711	14,678		3,050,553
Total	2,828,219	98.80%		1,181,602	96.38%		794,368	83.47%	
Assembly (Contigs)									
Sequences	83,758		30,999,497	138,906		36,661,374	92,208		24,775,032
N50			391			251			251
Putative viral seqs	2,127	2.54%		3,342	2.41%		1,982	2.15%	
Seqs without BLAST hit	78,125			131,932			88,346		
Assembly (Singlets)									
Sequences	714,161			501,137			113,903		
Putative viral seqs	18,107	2.53%		8,029	1.61%		5,487	4.81%	
Seqs without BLAST hit	517,133			413,810			38,781		
	Estimated Value		**SE**	**Estimated Value**		**SE**	**Estimated Value**		**SE**
Richness	831.2		27.2	843.5		37.9	865.6		43.8

Supplementary material 1. Metagenomic sequencing summary statistics (waste water seasonal samples). Sequences and nucleotide counts are total values, number of pairs is half of the shown values. Percent of reads sequence refer to the total amount of raw reads, while the percent of sequences having a BLAST hit, or not, corresponds to the total number of assembled sequences (contigs and singletons). Difference between the sequences assigned to known viruses and the sequences without a BLAST hit relates to those sequences having a BLAST hit that has not passed all the filtering criteria for a valid species assignment.

Supplementary material 3.

Table 1.- Distribution of adenovirus types identified in wastewater sample. The number of assigned OTU's, the number of 454 reads associated to those taxonomic units and the % of representative sequences are provided.

Adenoviridae member	Adenovirus type	GenBank Accesion number	OTU's	454 reads	% of sequences
Murine adenovirus 2		NC_014899	965	31440	59.41043084
Human adenovirus 41	HAdV-F	DQ315364	655	15473	29.23847317
Simian adenovirus 48		HQ241818	384	1238	2.339380197
Human adenovirus type 51	HAdV-D	DQ149642	248	864	1.632653061
Ovine adenovirus 6		DQ630759	268	677	1.279289494
Simian adenovirus 19		KP329565	158	500	0.944822373
Human adenovirus type 48	HAdV-D	EF153473	48	382	0.721844293
Human adenovirus 40	HAdV-F	L19443	119	342	0.646258503
Equine adenovirus 2		AEEHEXEND	103	318	0.600907029
Human adenovirus type 46	HAdV-D	AY875648	34	260	0.491307634
Simian adenovirus 47		FJ025929	112	250	0.472411187
Human adenovirus type 35	HAdV-B	AY271307S	73	175	0.330687831
Simian adenovirus 38		FJ025922	74	146	0.275888133
Bat adenovirus TJM		GU226970	53	143	0.270219199
Simian adenovirus B		KC693021	73	133	0.251322751
Human adenovirus 59	HAdV-D	JF799911	70	118	0.22297808
Simian adenovirus 18		FJ025931	43	82	0.154950869
Human adenovirus type 19	HAdV-D	DQ149618	23	63	0.119047619
Human adenovirus type 49	HAdV-D	DQ393829	14	57	0.107709751
Human adenovirus 68	HAdV-B	JN860678	8	36	0.068027211
Skunk adenovirus PB1		KP238322	21	28	0.052910053
Alpaca adenovirus		GQ499375	21	28	0.052910053
Simian adenovirus 16		NC_028105	8	21	0.03968254
Human adenovirus type 43	HAdV-D	DQ149636	2	16	0.030234316
Simian adenovirus 7		DQ792570	13	13	0.024565382
Simian adenovirus 42.3		FJ025925	10	13	0.024565382
Simian adenovirus 21		AC_000010	5	11	0.020786092
Human adenovirus type 1	HAdV-C	AF534906	4	9	0.017006803
Human adenovirus 54	HAdV-D	AB333801	2	9	0.017006803

Murine adenovirus 3		NC_012584	8	8	0.015117158
Simian adenovirus 11		KP329562	8	8	0.015117158
Simian adenovirus 43		FJ025900	4	6	0.011337868
Human adenovirus 12	HAdV-A	NC_001460	2	6	0.011337868
Simian adenovirus 13		NC_028103	5	5	0.009448224
Pigeon adenovirus 1		FN824512	4	5	0.009448224
Tree shrew adenovirus 1		AF258784	3	5	0.009448224
Human adenovirus type 31	HAdV-A	AM749299	4	4	0.007558579
Turkey adenovirus 1		NC_014564	1	4	0.007558579
Cynomolgus adenovirus		KT013209	3	3	0.005668934
Porcine adenovirus 5		AF289262	3	3	0.005668934
Simian adenovirus 40.2		FJ025926	2	3	0.005668934
Simian adenovirus 8		NC_028113	2	2	0.003779289
Human adenovirus 18	HAdV-A	GU191019	2	2	0.003779289
Human adenovirus 52	HAdV-G	DQ923122	2	2	0.003779289
Human adenovirus type 34	HAdV-B	AY737797	2	2	0.003779289
Murine adenovirus A		NC_000942	1	1	0.001889645
Simian adenovirus 46		FJ025930	1	1	0.001889645
Simian adenovirus 31.2		FJ025904	1	1	0.001889645
Fowl adenovirus 10		FAU26221	1	1	0.001889645
Human adenovirus type 20	HAdV-D	DQ149619	1	1	0.001889645
Human adenovirus type 7	HAdV-B	AC_000018	1	1	0.001889645
Porcine adenovirus 3		AB026117	1	1	0.001889645

Table 2.- Primers and conditions used for target enrichment assay for HAdV hexon are expressed in the following table:

PCR round	Gene (size)	Primer ID	Sequence (5'-3')	PCR Mix	Amplification conditions
1st	300	Forward	GCCSCARTGGKCNTACATGCACAT	In a final volume of 50 µl containing 1× Gold buffer (Applied Biosystems, Inc) at 50 mM, MgCl2 25 mM, 0,1 mM of each deoxynucleotide, 0,5 µM of each primer, 1U Taq polymerase and 10ul of the concentrate extraction	35 cycles of denaturation at 94ºC 30s, annealing at 50ºC 30s and extension at 72ºC 1min
		Reverse	CARNACVCCNCKRATGTCAAA		
2nd (introduction of barcodes & key index)	360	Forward	CCATCTCATCCCTG CGTGTCTCCGACTC AGGCCSCARTGGKC NTACATGCACAT	In a final volume of 50 µl containing 1× Gold buffer (Applied Biosystems, Inc) at 50 mM, MgCl2 25 mM, 0,1 mM of each deoxynucleotide, 0,5 µM of each primer, 1U Taq polymerase and 3ul of product from the first PCR	25 cycles of denaturation at 94ºC 30s, annealing at 50ºC 30s and extension at 72ºC 1min
		Reverse	CCTATCCCCTGTGTG CCTTGGCAGTCTCAG CARNACVCCNCKRATGTCAAA		

Samples	SMF1			SMF2			Ultra1			Ultra2		
	Sequences		Nucleotides	Sequences		Nucleotides	Sequences		Nucleotides	Sequences		Nucleotides
MiSeq Raw READS	3,846,130		962,076,715	4,079,546		1,039,169,432	4,078,152		1,015,816,890	3,441,666		924,822,940
Clean READS												
Pair-Ends	3,654,800		830,379,508	3,841,750		884,294,646	3,871,162		879,354,886	3,237,524		416,596,088
Single-Ends	176		40,654	998		238,833	170		40,190	136		32,184
Total	3,654,976	95.03%		3,842,748	94.20%		3,871,332	94.93%		3,237,660	94.07%	
Assembly (Contigs)												
Sequences	157,402		60,268,319	142,298		53,932,002	212,737		81,299,715	282,605		111,648,857
N50			391			386			387			396
Putative viral seqs	3,070	1.95%		2,017	1.42%		4,568	2.15%		5,629	1.99%	
Seqs without BLAST hit	147,950			135,870			199,878			265,817		
Assembly (Singlets)												
Sequences	3,372,435			3,509,087			3,466,437			2,663,042		
Putative viral seqs	18,647	0.55%		10,744	0.31%		24,825	0.71%		27,803	1.04%	
Seqs without BLAST hit	3,069,901			3,114,207			3,152,828			2,428,443		
	Estimated Value		SE	Estimated Value		SE	Estimated Value		SE	Estimated Value		SE
Richness		755.8	23.6		541	18.9		1066.4	23.2		1318.5	27.6

Supplementary material 4. Metagenomic sequencing summary statistics (concentration methods comparison). Sequences and nucleotide counts are total values, number of pairs is half of the shown values. Percent of reads sequence refer to the total number of raw reads, while the percent of sequences having a BLAST hit, or not, corresponds to the total number of assembled sequences (contigs and singletons). Difference between the sequences assigned to known viruses and the sequences without a BLAST hit relates to those sequences having a BLAST hit that has not passed all the filtering criteria for a valid species assignment.

ARTICLE IV

Avaluació amb tècniques de metagenòmica de la contaminació viral present a la superfície de plantes de julivert irrigades amb aigua de riu contaminada fecalment

4. Avaluació amb tècniques de metagenòmica de la contaminació viral present a la superfície de plantes de julivert irrigades amb aigua de riu contaminada fecalment.

Tenir accés a una quantitat d'aliments és una necessitat fonamental i un dret humà. Aquest dret s'estén al fet de consumir aliments microbiològicament segurs. Malgrat l'existència de normatives que vetllen per la qualitat dels aliments consumits, de forma continuada apareixen brots associats al consum d'aliments contaminats. Entre els principals agents etiològics responsables d'aquests brots alimentaris cal destacar els virus, i en particular els norovirus humans i els virus de l'hepatitis A i E.

L'aplicació de tècniques de seqüenciació massiva en aliments per a l'estudi de virus contaminants és un camp molt nou que ha estat poc explorat. En el present estudi s'han aplicat tècniques de seqüenciació massiva sobre un concentrat víric provinent de la superfície de plantes de julivert irrigades amb aigua de riu, que té un cabal molt depenent dels efluents de les plantes depuradores d'aigües que hi ha a la conca, i que per tant, conté contaminació fecal humana. Simultàniament també es va analitzar la presència de virus a l'aigua de reg utilitzada per regar els juliverts.

Importants espècies víriques de les famílies *Caliciviridae, Hepeviridae* i *Picornaviridae* van ser detectades a la superfície del vegetal. L'aigua de riu utilitzada per regar les plantes de julivert també va ser analitzada mitjançant tècniques de seqüenciació massiva, detectant-se membres de les famílies *Adenoviridae, Reoviridae*, *Picornaviridae i Astroviridae*. En aquestes mostres, també es van identificar seqüències relacionades de forma llunyana amb les famílies *Picornaviridae* i *Hepeviridae.*

En aquest treball s'ha presentat un protocol amb alta sensibilitat per a l'estudi de virus presents en aigua de reg i aliments, mitjançant tècniques de seqüenciació massiva. No obstant, la detecció de genomes vírics en una mostra no representa necessàriament l'existència d'una amenaça per a la salut. Es requereixen més estudis que permetin estimar la relació entre la presència de genomes amb el risc d'infecció. Les tècniques de metagenòmica, tot i permetre la detecció simultània de múltiples patògens en un únic assaig, necessiten ser millorades per al seu ús com a tècnica estandarditzada per l'anàlisi de virus en aliments, ja que els virus patògens solen trobar-se a molt baixa concentració en aquestes matrius. Aquest estudi també ha posat de rellevància problemes com la baixa proporció de seqüències assignades a virus o la inclinació de la tècnica a amplificar els genomes relativament més abundants, que haurien de ser solucionats en futurs estudis.

A metagenomic assessment of viral contamination on fresh parsley plants irrigated with fecally tainted river water

X.Fernandez-Cassi[1,¶, *], N. Timoneda[1,2¶], E. Gonzales-Gustavson[1], J.F. Abril[2], S.Bofill-Mas[1], R. Girones[1]

[1]Laboratory of Virus Contaminants of Water and Food, Department of Genetics, Microbiology and Statistics, University of Barcelona, Barcelona, Catalonia, Spain

[2] Computational Genomics Lab, Department of Genetics, Microbiology and Statistics, University of Barcelona; and Institute of Biomedicine (IBUB), University of Barcelona; Barcelona, Catalonia, Spain

[*] Corresponding author

[¶] These authors contributed equally to this work.

ABSTRACT:

Microbial food-borne diseases are still frequently reported despite the implementation of microbial quality legislation to improve food safety. Among all the microbial agents, viruses are the most important causative agents of food-borne outbreaks. The development and application of a new generation of sequencing techniques to test for viral contaminants in fresh produce is an unexplored field that allows for the study of the viral populations that might be transmitted by the fecal-oral route through the consumption of contaminated food. To advance this promising field, parsley was planted and grown under controlled conditions and irrigated using contaminated river water. Viruses polluting the irrigation water and the parsley leaves were studied by using metagenomics. To address possible contamination due to sample manipulation, library preparation, and other sources, parsley plants irrigated with nutritive

solution were used as a negative control. In parallel, viruses present in the river water used for plant irrigation were analyzed using the same methodology. It was possible to assign viral taxons from 2.4 to 74.88% of the total reads sequenced depending on the sample. Most of the viral reads detected in the river water were related to the plant viral families *Tymoviridae* (66.13%) and *Virgaviridae* (14.45%) and the phage viral families *Myoviridae* (5.70%), *Siphoviridae* (5.06%), and *Microviridae* (2.89%). Less than 1% of the viral reads were related to viral families that infect humans, including members of the *Adenoviridae*, *Reoviridae*, *Picornaviridae* and *Astroviridae* families. On the surface of the parsley plants, most of the viral reads that were detected were assigned to the *Dicistroviridae* family (41.52%). Sequences related to important viral pathogens, such as the hepatitis E virus, several picornaviruses from species A and B as well as human sapoviruses and GIV noroviruses were detected. The high diversity of viral sequences found in the parsley plants suggests that irrigation on fecally-tainted food may have a role in the transmission of a wide diversity of viral families. This finding reinforces the idea that the best way to avoid food-borne viral diseases is to introduce good field irrigation and production practices. New strains have been identified that are related to the *Picornaviridae* and distantly related to the *Hepeviridae* family. However, the detection of a viral genome alone does not necessarily indicate that there is a risk of infection or disease development. Thus, further investigation is crucial for correlating the detection of viral metagenomes in samples with the risk of infection. There is also an urgent need to develop new methods to improve the sensitivity of current Next Generation Sequencing (NGS) techniques in the food safety area.

Keywords: Food metagenomics, viral pathogens, food safety, *Caliciviridae, Picornaviridae*, next generation sequencing

1. Introduction

Food-borne diseases remain as a significant cause of illness worldwide, and consumers are exposed to microbiological and chemical contaminants. From a microbiological point of view, food can be a vehicle for protozoan, bacterial, viral, and prion infections. Although most fecally excreted microorganisms cause gastroenteritis or acute hepatitis, other pathologies such as meningitis, myocarditis, and neurological disorders are also possible.

Food contamination can occur at several stages of food chain production, from the irrigation and collection stages on farms to contamination during food processing in industrial settings, during food preparation at a restaurant, or at home. In high income countries, measures have been implemented to reduce the risk of fecal contamination of water and food such as proper sewer pipeline systems and hygienic measures during food handling, manufacturing, and preparation. Countries use legislation to control the microbiological quality of water and food, yet food-borne outbreaks are still reported frequently (Bernard et al., 2014; Ethelberg et al., 2010).

Among all the food-borne etiological agents, viruses are the most important causative agents of food-borne outbreaks (Painter et al., 2013). For example, noroviruses are the leading cause of food-borne illnesses, accounting for 125 million cases per year (Kirk et al., 2015). The increase in fresh food consumption, probably as a way for consumers to develop a healthier diet, has been linked to an increase in viral food-borne outbreaks (Callejón et al., 2015; Kozak et al., 2013). Coleman and co-workers (2013) noted that of the 127 outbreaks associated with leafy greens in the United States from 2004-2008 in the area where the pathogen was identified, 64% of cases were attributed to a viral infectious agent such as noroviruses, sapoviruses, and hepatitis A virus. As

mentioned, viruses can be accidentally introduced at different steps in food chain production, and crop irrigation with fecally contaminated water is one of the most critical points. Recently, Maunula et al. (2013) reported that 9.5% of irrigation water samples used to water berries were positive for human adenoviruses (HAdV), underlining the presence of fecal contamination. Although there are laws to control the microbiological water quality, none of them currently include specific viral parameters, and, therefore, water safety monitoring relies only on the use of fecal indicator bacteria (FIB). The usefulness of these laws for minimizing the viral presence in food matrices is unclear because the FIB do not always correlate with the presence of viral pathogens, and viruses are more resistant to water treatments than bacteria (Gerba et al., 1979; Jiang, 2006; Marzouk, 1980; Pusch et al., 2005; Savichtcheva and Okabe, 2006)

While bacterial contamination in food has been widely reported, notifications of viral outbreaks have been hampered in many cases by a lack of specific, sensible, and standardized concentration/detection methods. These important factors are mandatory for the inclusion of viral parameters in food legislation. Recently, a standard ISO 15216-1:2017 was published for the concentration and quantification of the two food-borne viruses hepatitis A virus (HAV) and human noroviruses (HNoV) (https://www.iso.org/standard/65681.html), and qualitative detection is in preparation (ISO/TS 15216-2:2013). The concentration methods are limited by their low recoveries while detection and quantification methods, which are usually based on RT-PCR or q(RT)PCR, are restricted to specific targeted viruses. Many different pathogens may contaminate water and food simultaneously, especially if fecally contaminated irrigation water is used on fresh vegetables. The introduction of next generation sequencing (NGS) techniques in the food safety field allows for the simultaneous analysis of

diverse viral pathogens in a single assay. To date, NGS has been applied to study the viral species that are present in all types of environmental and clinical samples as follows: oceans (Hurwitz et al., 2013), lakes (Djikeng et al., 2009), raw sewage (Cantalupo et al., 2011), reclaimed water (Rosario et al., 2009), and infectious clinical samples with unknown etiological agents (Greninger et al., 2015). However, the use of NGS in the food safety field has not been exhaustively explored (Aw et al., 2016; Park et al., 2011; Zhang et al., 2014).

The focus of the present study is to provide more information on the applicability of NGS techniques in the food safety field by studying viral contamination in fresh vegetables, with parsley plants (*Petroselinum crispum L*) that were irrigated with river water containing fecal contamination as the model.

2. Materials and methods

2.1. River water samples

The Besòs River (Sant Adrià de Besòs, Barcelona, Catalonia, Spain) is a river that is 17.7 km long, and it displays irregular discharge due to the Mediterranean climate. Along its path, the river collects the secondary effluent of 27 wastewater treatment plants (WWTP) and ends in the Mediterranean Sea next to Barcelona. Since the mid-1990, it has been subject to a recovery process to improve its water quality. As a consequence of these efforts, the Besòs River is being used as a source of water to irrigate crops by some local farmers. Twenty liter river water samples from the Besòs River were collected in May 2014. Ten L of each river sample was used to irrigate the parsley plants, and the remaining 10 L was used to concentrate the viral particles to characterize the viruses that were present through qPCR and NGS. After two weeks, the procedure was repeated and the

river water was tested by qPCR. The river water samples collected for parsley irrigation were kept in the dark at 4°C for 15 days and used to irrigate the plants.

2.2. River water viral particle concentration

The viruses that were present in the 10 L river water samples were concentrated using skimmed milk organic flocculation. This method has a recovery efficiency of 50% (20–95%) and it was applied as previously described by Calgua et al. (2013). In brief, a suspension of skimmed milk was prepared by adding 10 g of skimmed milk powder (Difco, Detroit, MI, USA) to 1 L of artificial seawater (Sigma-Aldrich Chemie GMBH, Steinheim, Germany); the solution was then adjusted to a pH of 3.5 using 1 N HCl to obtain a pre-flocculated 1% (w/v) skimmed milk solution (PSM). At that moment, all the river water samples were adjusted to a conductivity of 1.5 mS/cm^2 and acidified to a pH of 3.5 using 1 N HCl. Ten mL of PSM was added to previously conditioned samples to obtain a final sample milk concentration of 0.01%. After 8 h of stirring at room temperature, the flocs were allowed to sediment by gravity for 8 h. The supernatant was carefully removed, and the remaining sediment from approximately 500 mL of the solution was centrifuged at 8000 × g for 30 min at 4°C. The pellets were suspended in 8 mL of phosphate buffer (pH 7.5) and stored at-80°C until nucleic acid (NA) extractions could be performed. The two viral concentrates are labeled as the Besòs River Water (BRW1 and BRW2) samples.

2.3. Parsley plant growth and irrigation

Parsley (*Petroselinum crispum L)* seeds were planted and cultivated in a climate room at 22°C, 60% relative humidity, and light conditions equivalent to 110 µmol

of photosynthetically active radiation (PAR) at the "Serveis de Camps Experimentals" at the University of Barcelona. The seeds were irrigated using a nutritive solution containing iron, nitrogen and phosphorus (Hoagland solution 50%). After 6 weeks, a total of 12 different parsley pots were moved to a greenhouse from the same facility.

All 12 parsley pots were irrigated twice a week during the whole growth process (from May to June) by using 4 L of the same nutritive Hoagland solution. This irrigation procedure was performed by inundating the tray containing the parsley pots, and thus the parsley leaves were not washed. Half of the 12 pots were irrigated daily, in addition to receiving nutritive solution, by spraying the leaves with 15 mL of Besòs river water. These samples were labeled as Besòs River Parsley (BRP).

In the same way, the remaining pots were used as a Negative Control Parsley group (NCP), and they were irrigated daily by spraying 15 mL of the same nutritive solution that was used for irrigation, but by inundation, into both the control and test pots. The negative control was used as a blank sample to identify viral sequences that could be naturally present in the parsley plants or due to other external factors (greenhouse, irrigation nutritive solution, facility users, manipulation, reagent contaminants or equipment).

Two weeks later, a fresh BRW sample was collected to continue the irrigation process for another 15 days. At the end of the study, both plant groups were irrigated with 450 mL of river water or control nutritive solution water. After one month of daily irrigation, 25 g of parsley leaves from NCP and BRP were hand-cut by investigators who were wearing sterile gloves, and the leaves were placed in sterile bags (BagPage® filter-bag, Interscience, France). The samples were kept at 4°C for less than 48 h until the concentration method was applied.

2.4. Viral concentration from plants

Twenty-five grams of NCP and BRP were washed in a sterile filter-bag with 50 mL of glycine buffer (pH of 9.5, 0.25 N) for 40 minutes using a stomacher. Afterwards, the sample pH was adjusted to 7.0 (±0.2) by using HCl 0.1 N. To remove the bacteria and other suspended organic material, the samples were centrifuged at 8000 x g for 10 minutes at 4°C. The supernatant was carefully collected without disturbing the pellet and ultracentrifuged at 90,000 x g for 1 h. The supernatant was discarded and the pellet was suspended in a final volume of 500 µl of phosphate buffer (pH 7.5). Viral concentrates were stored at-80°C until further analysis.

2.5. *Nucleic acid extraction, library preparation, and sequencing*

Volumes of NCP, BRP, BRW1 and BRW2 viral concentrates were treated with 156 units of Turbo DNAse (cat nº AM1907, Ambion,Lithuania) for 1 h at 37°C to remove free DNA prior to nucleic acid extraction, and 280 µl of the treated DNAse viral concentrate was extracted using the QIAamp®Viral RNA Mini Kit from QIAGEN (Qiagen, Valencia, CA, USA). The nucleic acids (NA) were eluted in 60 µl according to the manufacturer's instructions, and they were stored at-80 °C for further analysis. To enable the detection of both DNA and RNA viruses, the total NAs were reverse-transcribed as previously described in Wang et al. (2002 and 2003). In short, SuperScript II (Life Technologies, California, USA) was used to retrotranscribe RNA to cDNA with primerA (5'-GTTTCCCAGTCACGATCNNNNNNNNN-3'). Second strand cDNA and DNA were constructed with the primer sequences using Sequenase 2.0 (USB/Affymetrix, Cleveland, OH, USA). A PCR amplification with AmpliTaqGold (Life Technologies,

Austin, Texax, USA) was performed using primerB (5'-GTTTCCCAGTCACGATC-3') with 20-30 cycles; this step was run in duplicate. The PCR products were purified and eluted in 15 µl using a Zymo DNA clean and concentrator (cat nº D4013, Zymo Research, USA) to yield enough DNA for the library preparation. NGS sequencing was performed by SGB-UAB, Barcelona. Amplified DNA samples were quantified by Qubit 2.0 (Life Technologies, Oregon, USA) and libraries were constructed using a Nextera XT DNA sample preparation kit (Illumina Inc) in accordance with the manufacturer's instructions. The samples were sequenced on an Illumina MiSeq 2x300 in base-pair paired end format. Raw read sequences for the four analysed samples were made publicly available at NCBI Sequence Read Archive database (NCBI-SRA), under the PRJNA381682 bioproject.

2.6. *Determination of the level of human fecal contamination in river water by HAdV qPCR*

To evaluate the level of human fecal contamination in the Besòs river water, specific real-time qPCR assays for human adenovirus were performed with TaqMan® Environmental Master Mix 2.0 (Life Technologies, Foster City, CA, USA). Real-time primers and probes for HAdV were described in (Bofill-Mas et al., 2006; Hernroth et al., 2002). An MX3000Pro qPCR sequence detector system (Stratagene, La Jolla, CA, USA) was used to quantify the HAdV.

2.7. *Bioinformatic analyses*

The quality of the raw and clean read sequences was assessed using FASTX-Toolkit software, version 0.0.14 (Hannon Lab, http://www.hannonlab.org). The read

sequences were cleaned by Trimmomatic version 0.32 (Bolger et al., 2014), while attending to the sequencing adaptors and linker contamination. Low quality ends were trimmed by using an average threshold Phred score above Q15 over a running-window of 4 nucleotides. Low complexity sequences, which were mostly biased to repetitive sequences that would affect the performance of downstream procedures in the computational protocol, were then discarded after estimating a linear model based on Trifonov's linguistic complexity (Sarma et al., 1990) and the sequence string compression ratio. The discrimination criteria for that linear model assumes low complexity scores below a line with a 45º slope and it crosses at 5% below the complexity inflexion point found by the model, which is specific to each sequence set. Finally, duplicated reads were removed in a subsequent step to speed up the downstream assembly.

Assemblies were obtained using Velvet (Zerbino and Birney, 2008) and Meta-Velvet tools (Namiki et al., 2012) versions 1.2.10 and 1.2.02, respectively. Afterwards, the contigs and singletons longer than 100 bp were queried for sequence similarity using NCBI-BLASTN and NCBI-BLASTX (Altschul et al., 1997, 1990) against the NCBI viral complete genome database (Brister et al. 2015), the viral division from the GenBank nucleotide database (Benson et al., 2015), and viral proteins from UniProt (UniProt Consortium 2015, ftp://ftp.uniprot.org/pub/databases/uniprot/current_release). The species nomenclature and classification were assigned according to the NCBI Taxonomy database standards and the basic Baltimore classification. The HSPs considered for taxonomic assessment must have an E-value that is lower than 10^{-5} and a minimum length of 100 bp. On the basis of the best BLAST results and a 90% coverage cutoff, the sequences were classified into their likely taxonomic groups of origin. Contigs and singletons with viral taxonomic assignments were

scaffolded by using the Geneious software assembler (Geneious 9, Kearse et al. 2012). The scaffolded sequences were subsequently mapped using the Geneious mapper tool. Phylogenetic trees were constructed, also using Geneious software and a neighbor-joining method with 1,000 bootstrap replicates was chosen.

2.8. *Richness*

Tables summarizing the number of sequences from the assembly matching of each taxonomic unit were built. From those tables, the richness ratios were calculated with CatchAll software, version 4.0 (Allen et al., 2013). Among all the models included in the package, the Chao1 non-parametric model was chosen, which was the one that provided the best results about the datasets.

3. *Results*

3.1. Mi-Seq sequence output

Parsley that was irrigated with river water was assayed using 20 (BRP20) and 30 amplification cycles (BRP30). The 20 and 30 river irrigated parsley samples, the parsley negative control and the river water samples (BRP20, BRP30, NCP, BRW1 and BRW2) were sequenced in a Mi-Seq Illumina device, producing 1,192,716, 6,677,622, 8,067,728, 5,426,788 and 2,476,054 raw reads, respectively. A summary of the bioinformatics filtering and assembly of the reads can be found on Table 1 (data for BRP30 not shown, as explained on next section). Total number of sequences having at least an homology hit to a viral sequence from the distinct databases after BLAST searches, as well as the total viral species and families from which each sample richness score was estimated are also listed on that table. Only a small amount of sequences were taxonomically assigned to known

viruses, except for the NCP sample where most of the hits accounted for viruses infecting aphids parasiting the leaves. In this case, despite having much more reads from the sequencing experiment, the taxonomic complexity was lower anyway.

3.2.*Estimating the viral diversity by using CatchAll*

The estimated total richness ratios for the BRP20 and BRP30 samples were 255.6 (±34.7 s.d.) and 171.8 (±40.9 s.d.), respectively. The higher estimated viral richness obtained in BRP20 (20 cycle amplification) over BRP30 (30 cycles amplification) is consistent with the observed higher diversity of the detected viral species. It has been described that number of cycles can introduce biases on sequence representation as an effect of the exponential amplification of the PCR over content specific nucleotide sequences (Kebschull and Zador, 2015). For this reason, all further analyses were conducted only taking into account BRP20.

The resulting richness ratios for NCP, BRW1, and BRW2 were 130.3 (±34.7 s.d.) and 632.4 (±44.4 s.d.) and 923.5 (±70.9 s.d.), respectively. The richness value obtained from the river water is about 3 to 4 times greater than that obtained for the other two samples. This finding reflects a higher diversity in the viral population that was present in a more complex ecosystem, such as river water, which is heavily impacted by human activity. As expected, the NCP richness was lower than that of BRP, which was irrigated using BRW.

3.3. Identification of human and animal viruses in river water samples used for irrigation (BRW)

River water samples from the Besòs River contained human viral fecal contamination, as shown by the finding that both samples were positive by qPCR for Human Adenovirus (HAdV) with low concentration $1.2x10^{+3}$ GC/L and $5.62x10^{+2}$ GC/L, respectively. Parsley irrigated with Bèsos River water was qPCR-positive for HAdV, also with low concentration $7.2x10^{+1}$ CG/25g. Interestingly, only the BRW2 presented reads assigned to *Adenoviridae* family, specifically to HAdV41 (HAdV-F). Plant-infecting ssRNA+ viruses from the *Tymoviridae*, *Virgaviridae*, and *Alphaflexiviridae* families were found. These plant viral families accounted for a 66.13%, 14.45%, and 1.11% of the viral reads, respectively. The primary phage sequences found in the river water were annotated as *Myoviridae*, *Siphoviridae, Podoviridae*, and dsDNA *Microviridae* families, accounting for 5.70%, 5.06%, 3.69% and 2.88% of the viral reads, respectively. The human and potentially zoonotic viruses found here are summarized in Table 2. By applying the described metagenomics approach, up to 26 different viral families were identified in river water used for irrigation. Among the pathogenic and potentially pathogenic viruses, members of the *Astroviridae*, *Adenoviridae, Reoviridae, Picobirnaviridae*, *Picornaviridae* and *Parvoviridae* families were detected. Although important pathogenic viruses were detected using the metagenomics procedure, the sum of all the reads from all the putative pathogenic families did not reach the 1% of viral reads for the sample.

3.3.1 ssRNA+ viruses detected from Picornaviridae, Astroviridae and Hepeviridae families in BRW

Human pathogenic viruses within the *Picornaviridae* family, such as Aichi virus species, were found to have high identities (94.4%). Compared to reference strains several sequences presented low identity scores and high coverage (>99%); been distantly related to the recently described *Ampivirus* genus (NC_027214.1) and were detected in the BRW1. The BLASTx sequences matched up to 80% of the identity of putative RHV-like sequences in the ampivirus species of the *Picornaviridae* family, suggesting that these detected viral sequences could represent a new viral species within the same genus. Furthermore, sequences of 159 and 206 base pairs resembling rodent hepatovirus and bat picornavirus were detected, revealing shared identities of 77 and 79%, respectively.

Several sequences that were distantly related to the *Hepeviridae* family were detected in Besòs river water. A total of 4 sequences with lengths ranging from 242 to 636 nucleotides showed a distant relation to a sequence from a newly suggested virus named tentatively hepelivirus (accession AFR11847) according to a BLASTx against the entire GenBank database. All the translated reads were distantly aligned to the RdRp region of this new proposed virus with identities ranging from 39.3 to 42.2% and query coverage from 54.43 to 100%. A translated sequence from the read with query coverage of 100% was used to construct a phylogenetic tree, together with reference to *Hepeviridae* GenBank entries covering the RdRp region (Figure 1). Within the *Hepeviridae* family, the sequence detected in river water shares its closest common ancestor with Hepelivirus, which was recently described in a sewage sample from Nepal. Both sequences seem to have diverged from species within the *Orthohepevirus* genus and are closer to the recently characterized *Piscihepevirus*. Sequences belonging to the *Astroviridae* species Mamastrovirus 1 were detected in river water that was used

for irrigation, with nucleotide identities ranging from 92.6% to 97.4%. Most of the resulting sequence reads had a hit with the serine protease that is encoded in the genome of this viral family, showing high identities at the amino acid level (≥90%).

3.3.2 ssDNA viruses detected in from families Parvoviridae and Circoviridae in BRW

The ssDNA viral reads assigned to the *Parvoviridae* family represented 0.067% of the total viral sequences. Most of the viral reads associated with this family are sequences that share a high amount of their identity with several densoviruses. Those species infect a wide range of hosts including cockroaches, mosquitoes, and ants, among other. However, some of the sequences presented a heterogeneous degree of identity with viruses that infect mammals (swine, rats, and mice) and with several adeno-associated parvoviruses that infect humans (84-100%). ssDNA viral reads matching the *Circoviridae* family made up 0.07% of the viral sequences. Most of the reads associated with this family presented a high identity with several circoviral sequences that were previously described in porcine stool, birds, insects and sewage.

3.3.3. dsRNA viruses of Reoviridae, Picobirnaviridae families detected in BRW

Human Rotaviruses from species A were detected only in BRW2 sharing higher identities (99-100%) whereas BRW1 had no reads taxonomically assigned to the family *Reoviridae.*

Sequences that were remotely homologous to the double-stranded RNA human picobirnaviruses were detected. These sequences displayed proteins similar to

the RNA-dependant RNA polymerase (RdRp) region of the human picobirnaviruses, showing identities ranging from 60 to 90%. Some reads hit the capsid viral protein and ORF3 of human picobirnaviruses.

3.4. *Identification of human and animal viruses in parsley plants irrigated with control water (NCP)*

NCP was used as a negative control to discard the sequences that were already present in the plants or introduced due to external factors. Despite the higher number of viral sequences obtained (7.50%) in comparison with the sequences from BRW (2.41%), none of them matched the pathogenic enteric viruses. Most of the taxonomically assigned sequences from this sample were classified as the dsDNA phage families *Podoviridae* (32.09%) and *Siphoviridae* (0.09%), whereas the majority of ssRNA+ viral sequences (57%) were related to the insect virus *Dicistroviridae* family.

3.5. Identification of human and animal viruses in parsley plants irrigated with river water samples (BRP)

Viral concentrates from the BRP contained viruses from 18 different families. Despite the low MiSeq output (596,358 raw reads), members of the families *Astroviridae, Caliciviridae, Flaviviridae, Hepeviridae, Parvoviridae*, and *Picornaviridae,* which included several viral pathogenic species, were detected. The *Dicistroviridae* family of insect viruses accounts for the majority of reads (41.52%). Among bacteriophages, the *Podoviridae* family are the most abundant, accounting for 14.06% of the total viral reads. Candidate sequences for the plant viral families *Tymoviridae* (0.19%) and *Virgaviridae* (0.13%) as well as the viral

phages families *Inoviridae* (1.77%) and *Siphoviridae* (0.19%) were also found. A more detailed list of human viral families is summarized in Table 3.

3.5.1. *ssRNA+ viruses from the Picornaviridae, Caliciviridae, Hepeviridae and Flaviviridae families detected in BRP*

The majority of human pathogenic viral reads from BRP were assigned to the *Picornaviridae* family (0.035% total reads from sample) and several enteroviruses from species A and B were detected. For enterovirus species A, most of the reads were similar to the reads for coxsackievirus A6 (CV-A6), coxsackievirus A10 (CV-A10), and coxsackievirus A16 (CV-A16). After scaffolding almost the entire genome, CV-A16 was recovered. Enterovirus species B reads displayed matches up with echovirus E6 and enteroviruses B4 and B5.

Viral sequences related to the ssRNA+ *Caliciviridae* family were identified and taxonomically assigned to human sapovirus GI and norovirus genogroup IV. The sapovirus reads were scaffolded with Geneious software. A contig of 1592 base pairs covering the RdRp/VP1 junction was found to be closely related to GI.2 sapoviruses, sharing 96% of the identity of GenBank accession AB614356.

Four sequences presented identities above 98% with norovirus GIV. One out of 4 NoVGIV sequences recovered here shared 100% of its nucleotide identity with the VP1 region (GenBank accession LC150859).

Within the *Hepeviridae* family, several sequences related to hepatitis E genotype 3 ranging from 154 to 548 base pairs long were detected in the parsley leaves. Those sequences matched up with variable identities at nucleotide levels ranging from 86% to 98%. A phylogenetic tree including the proposed HEV reference strains in Smith et al. (2016) was constructed. The HEV contig was sequenced from parsley leaves clusters and shown to have HEV genotype III sub genotype f

(data not shown). A sequence of 448 nucleotides that were taxonomically assigned to this family was detected by BLASTx. This sequence had a low amino acid identity match with the RdRP domain from the Hepatitis E virus (34% amino acid identity, 80% coverage, Accession ANJ02846). The retrieved sequence contains an FKGDDS domain as well, which is commonly present in other *Hepeviridae* family members. The sequence was aligned with other representative sequences within the *Hepeviridae* family, including the recently described strain Hepelivirus a sequence from sewage collected in Nepal. The phylogenetic tree suggests that the RdRp region from this putative *Hepeviridae* sequence that was detected on the parsley surface would be closer to the Hepelivirus sequence detected in sewage (Figure 1).

The ssRNA+ viral species GBV-C virus from the *Flaviviridae* family were also reported. These detected sequences had an identity between 84.7 and 94% at the nucleotide level, and they covered several regions of the GBV-C virus including the E1 and E2 proteins, helicase, NS3, and NS5A domains.

3.5.2. Members of ssDNA viral families Parvoviridae, Anelloviridae and Circoviridae detected in BRP

The ssDNA viral reads belonging to the *Parvoviridae* family represented 0.016% of the total viral sequences. Most of the reads were associated with human bocavirus. A large contig of 1,287 base pairs aligned with human bocavirus 2 (HBoV-2) (99.1% identity, 100% coverage). This region matches more than half of ORF3, a region that encodes the VP1 and VP2 viral capsid proteins. Only two reads of 99 and 451 base pairs match the densovirus.

Sequences related to the ssDNA viral *Anelloviridae* family are highly prevalent,

accounting for 35.10% of the total reads from the samples. A variety of sequences related to Torque teno mini virus (from TTMV 1 to 9), Torque teno midi virus (from TTMDV 1 and 2) and Torque teno virus (TTV2, 3, 18 and 19) were found. The most prevalent species were TTMVs 4, 5 and 8.

4. Discussion

In this manuscript, NGS techniques were applied using viral Metagenomics within the food safety field. River water contaminated with human faeces was used to irrigate parsley as a plant model. To our knowledge, only two studies using NGS techniques in the field of food safety have been published, with one on meat and another one on lettuce (Aw et al., 2016; Zhang et al., 2014), and many questions and technical improvement are still pending. In this study, a high diversity of viruses including pathogens has been identified in water and more importantly on the surface of vegetables. Contaminated parsley could present a risk to the population, and it could represent a source for different viral diseases such as hepatitis and gastroenteritis due to the presence of viral genomes from hepatitis E virus, sapoviruses, and several enteroviruses that are transmissible by the fecal-oral route.

BRW contains human viral fecal contamination as shown by the positive HAdV qPCR results and the results obtained by NGS. HAdV has been widely used as a fecal marker of human viruses as reviewed in Bofill-Mas et al. (2013). Therefore, the Besòs river water was appropriate for this simulation assay as representative irrigation water with human viral fecal contamination that could be used for crop irrigation. Despite the positive HAdV qPCR results, although with low concentration of adenovirus, for the two river water types collected here ($1.2x10^{+3}$ GC/L and $5.62x10^{+2}$ GC/L) and at the parsley surface ($7.2x10^{1}$ GC/25g), only HAdV41 sequences were detected in BRW2 by using untargeted

510 metagenomics. Viral metagenomics is usually limited by the quantities of
511 DNA/RNA of viral origin that are present in the tested samples. Therefore, after
512 the RT and sequenase reactions, a PCR-based random amplification method
513 (SISPA) is used to obtain enough DNA for library preparation. This PCR step might
514 introduce bias by amplifying the most abundant genomes and producing %GC
515 bias (Duhaime et al., 2012). Less abundant genomes might not be sequenced or
516 may be underrepresented (Karlsson et al., 2013). A balance between the number
517 of cycle amplifications and DNA concentrations for library preparation must be
518 achieved. The application of a higher number of PCR cycles produced higher DNA
519 concentrations for library preparation but decreased the estimated viral richness
520 of the river water-irrigated parsley, as in BRP30(data not shown). For these
521 reasons, metagenomes cannot be interpreted quantitatively but only as a relative
522 abundance of species for comparison. The detection of low numbers of DNA
523 viruses, such as Human Adenovirus, whose presence has been quantified in River
524 water and Parsley leaves by qPCR and have only been detected by metagenomics
525 in one river water sample, arises a sensibility question regarding metagenomics
526 applied to food safety. This point is especially important due to the fact that
527 viruses are present in food and environmental samples In low concentration,
528 many times close to the LODs and LOQs of PCR and (RT)qPCR systems to detect
529 them. To reduce this possible lack of sensitivity of viral metagenomics
530 targeted/capture sequence NGS could be a more suitable approach to be applied
531 as a tool to detect all diverse viral strains in one specific group of viruses in food
532 and drinking water. As sequencing costs are dramatically diminishing with the
533 novel high-throughput technologies, this kind of approaches will be affordable to
534 run periodic controls and to check potential risks in real time in the near future.
535 Noroviruses from GI and GII are abundantly found in fecally contaminated water
536 but were neither detected in vegetal produce nor in the river water tested. The

absence of these pathogens can be explained by their seasonality (Haramoto et al., 2006; Nordgren et al., 2009) and, probably, by the low titer of these viruses during the period of the year when the study was conducted (May). Other dsRNA viruses from the family *Picobirnaviridae* were detected in tested samples. Therefore, according to our results, there is no data that clearly indicate a bias for rotaviruses or other viruses not detected, and the absence of these virus might be related to a lower relative abundance compared to other viral species present.

Other factors that might affect the detected viral sequence diversity can be the different efficiencies of the given concentration methods and the representative sample volume tested for each matrix as well as the inactivation and degradation of viral genomes during crop production.

Despite this, the results presented in this work support the use of a viral Metagenomics approach for a general assessment of viral contaminants present in food matrices by using a single assay. Taking into account all the samples, viruses were assigned to more than 34 different families. Important viral pathogens belonging to the families *Astroviridae, Adenoviridae, Reoviridae, Caliciviridae*, *Circoviridae, Hepeviridae, Picornaviridae*, and *Parvoviridae* have been detected.

The parsley that was irrigated with river water shows a very high diversity of viral strains, and more pathogenic viruses were detected in parsley than in river water. This finding could be related to a lower background of DNA/RNA input in the sample and potentially high absorption or stability capabilities for some viral species at the vegetable surface. Two different concentration protocols have been used to concentrate viral particles due to the different nature of matrices tested. The viral concentration method used for parsley leaves is based on elution using glycine at basic pH and ultracentrifugation whereas the concentration from

564 river samples was done by using SMF. Although the recovery efficiency of both
565 methods have showed to be equivalent for viral indicators, the impact of different
566 concentration and extraction methods might have played a role as proven to
567 occur in sewage on the context of metagenomics sequencing (Hjelmsø et al.,
568 2017). Parsley that was irrigated with a nutritive solution and used as a negative
569 control showed an expected absence of viruses associated with fecal
570 contamination.
571 Human astrovirus (HAstV) is suspected to be involved in 0.5 to 15% of all acute
572 diarrhea outbreaks in children, representing the third-most common cause of
573 this pathology within that age segment (Bosch et al., 2014). In the present study,
574 all the HAstV-related sequences were only detected in BRW and were assigned
575 to the genus *Mamastrovirus*, species MAstV-1. The involvement of HAstV in
576 foodborne outbreaks has also been documented (Oishi et al., 1994), but its
577 occurrence seems to be of lower importance compared to the number of
578 outbreaks caused by NoV. As an example, the CDC's Foodborne Disease Outbreak
579 Surveillance System database (FOOD Tool -
580 https://wwwn.cdc.gov/foodborneoutbreaks) has reported only 2 outbreaks
581 involving Astroviruses since 1998. The importance of HAstV might be
582 underestimated due that in sporadic gastroenteritis cases the etiological agent is
583 not investigated.
584 The *Caliciviridae* viral family has been detected in the parsley plants. This family
585 contains 5 different genera including *Sapovirus* and *Norovirus,* and they are
586 known as human caliciviruses. Noroviruses (NoV) are the leading cause of food-
587 borne disease outbreaks worldwide (Koo et al., 2010) whereas human
588 sapoviruses (HSaV) are increasingly recognized as a food-borne outbreak
589 etiological agent (Kobayashi et al., 2012; Yamashita et al., 2010); it became the
590 third-most common cause of viral gastroenteritis after NoVs and rotavirus in

591 Osaka, Japan (Iritani et al., 2016). Human gastroenteritis is primarily caused by
592 the specific genotype NoV GII.4 (Verhoef et al., 2015). Other minority HNoV
593 species, such as GIV NoV (which are not commonly included in environmental or
594 food safety studies) were detected in the parsley leaves. NoV GIV food-borne
595 outbreaks have rarely been reported, but the association of this virus with human
596 pathology has been shown (Ao et al., 2014). The availability of environmental
597 NoV GIV information is scarce; therefore, information about the seasonality of
598 the virus is not clear. The presence of norovirus in fresh vegetables is a matter
599 of concern with respect to the norovirus genogroup, and more data regarding
600 NoV GIV seasonality and epidemiology is needed.
601 Human sapoviruses (HSaV) belonging to GI, GII, and GV were detected. Most of
602 the sequences were assigned to HSaV GI.2, which is considered a minor
603 genogroup with very few outbreaks reported (Iwakiri et al., 2009). However, an
604 increase in the number of outbreaks ligated to HSaV GI.2 has been observed in
605 recent years (Lee et al., 2012; Svraka et al., 2010;).
606 Interesting results have been obtained for the *Hepeviridae* family. Within the
607 *Orthohepevirus* genus, there are 4 classical Hepatitis E virus (HEV) genotypes that
608 can infect several avian and mammal hosts. Genotypes 1 and 2 have only been
609 detected in humans while genotypes 3 and 4 are considered zoonotic, being that
610 pigs, wild boars, and deer are the main reservoirs (Legrand-Abravanel et al.,
611 2009). HEV is an emerging virus that is transmitted through the fecal-oral route
612 and causes hepatitis. To our knowledge, HEV Food-borne outbreaks have been
613 linked to the consumption of contaminated meat (Li et al., 2005; Yazaki et al.,
614 2003) but no single outbreak has been linked to the consumption of fresh
615 vegetables yet. However, Kokkinos et al. (2012), who studied the prevalence of
616 HEV in leafy green irrigation water and in point-of-sale lettuce, have found the
617 virus in 1 out of 20 (5,0%) and 4 out of 125 (3,2%) of the tested samples,

618 respectively. HEV-3 has been detected in the parsley; this is a potentially zoonotic
619 HEV strain, and its finding is consistent with the results of Kokkinos et al., which
620 suggests that vegetables could represent a potential vehicle of transmission for
621 HEV. In Spain, HEV-3 in the clade of subtype 3f strains have been detected in most
622 autochthonous human cases (Fogeda et al., 2009) and in pigs, including HEV
623 cases associated to pork meat consumption (Riveiro-Barciela et al., 2015). In this
624 study, several sequences that are distantly related to the RdRp from genus
625 *orthohepevirus* and closely related to hepeliviruses found in sewage from Nepal
626 have been detected (Ng et al., 2012). This result shows how little actual
627 knowledge there is about the *Hepeviridae* family, but it will surely be expanded
628 with the application of NGS techniques to different samples. However, the role
629 and importance of these viral sequences as a food-borne agent have yet to be
630 fully understood.
631 Different viruses in the *Picornaviridae* family have been detected in river water
632 and parsley. *Picornaviridae* is a family grouping of more than 30 different genera
633 of ssRNA+ viruses that infect vertebrates and includes historically important
634 human pathogens, such as hepatitis A virus (HAV), enteroviruses (EV), and
635 poliovirus. They are very stable in the environment, and several waterborne and
636 foodborne outbreaks have been reported in this taxon (Severi et al., 2015). Aichi
637 virus (AiV), which has been detected in river water, is a viral specie within genus
638 *Kobuvirus* related to gastroenteritis (Yamashita et al., 1991). Recent studies have
639 shown that AiV may co-infect with other enteric viruses (Räsänen et al., 2010).
640 Different *Enteroviruses* (EV) from species A and B were detected in parsley
641 surface. An increase in enterovirus outbreaks has been reported recently by
642 emerging recombinant EV strains (Holm-Hansen et al., 2016; Zhang et al., 2010).
643 Nearly the entire genome of CV-A16, an emerging enterovirus genotype from
644 species A which is related to hand, foot, and mouth disease (HFMD), was

645 sequenced. Although EV can be transmitted person-to-person, the EV from these
646 species have been linked to water-borne (Beller et al., 1997; Häfliger et al., 2000)
647 and food-borne outbreaks (Le Guyader et al., 2008).
648 Closely related animal *Picornaviruses* sequences with low nucleotide identity to
649 *Ampivirus* (Reuter et al., 2015), a new genus infecting amphibians, and with rat
650 hepatovirus (Drexler et al., 2015), a proposed ancestral virus with a putative
651 common origin with HAV, have been detected.
652 Currently, several new picornaviruses have been discovered with the introduction
653 of NGS technologies to virology (Holtz et al., 2009; Ng et al., 2015). The low
654 shared identities for the sequences found in this study suggest that more
655 members of the *Picornaviridae* family will be discovered in the years to come.
656 Other viral sequences detected in the parsley or irrigation water used and whose
657 transmission route through fecally contaminated water or food is not fully
658 understood were detected. Those families include the GBV-C virus from the
659 *Flaviviridae* family, sequences remotely close to Human Picobirnaviruses, and
660 several ssDNA viruses from the families *Parvoviridae, Circoviridae and*
661 *Anelloviridae*. Those viruses are known to infect human although their
662 association with disease is unknown.
663 Several sequences that were taxonomically assigned to *Dicsitroviridae* family
664 aligned with the Aphid lethal paralysis virus (ALPV). The presence of greenflies in
665 the greenhouse area could explain their high abundance. Plant viruses from
666 *Virgaviridae* and *Tymoviridae* families as well as bacteriophages from the
667 *Siphoviridae*, *Podoviridae*, *Microviridae* and *Inoviridae* families were detected.
668 The quantities of plant and bacteriophage viruses matching our read sequences
669 were expected because previous environmental studies already highlighted their
670 high concentrations (Cantalupo et al., 2011; Ng et al., 2012).
671 Recently, some authors have described the ability of lettuce to internalize viral

672 particles belonging to the *Caliciviridae* family (DiCaprio et al., 2015a, 2015b;
673 Esseili et al., 2012b), which suggests the possibility that infectious viruses may
674 not only be found on food surfaces. Also the risk of infection through the
675 consumption of internalized NoV has been quantified using QMRA approaches
676 (Sales-Ortells et al., 2015). Present study has only evaluated viruses attached at
677 leaves surfaces. Future studies should gaze at this internalization scenario to
678 evaluate viral risk of infection through internalized viruses.
679 Finally, it is remarkable that the Metagenomics approach facilitates the capture
680 and assignment of a wide diversity of DNA/RNA viruses from a particular sample
681 in a single assay. In the present work, several viral species, the transmission
682 routes of which are not yet fully understood, have been detected in food or in
683 river water used for irrigation. Viral persistence has been associated with specific
684 and non-specific attachment to carbohydrate moieties (Esseili et al., 2012a). This
685 association has been described for human NoV VLPs, which are bound with
686 different strengths to the extracts of different plants including coriander, iceberg
687 lettuce, spinach, or romaine lettuce (DiCaprio et al., 2015b). Data involving
688 parsley extracts or extending to other viral models that were different from
689 caliciviruses have not been published to date.
690 Once viral genomic sequences have been detected, the question whether those
691 viruses represent a risk arises. The sole presence of viral genomes in food does
692 not necessarily represent a biological hazard, as natural inactivation processes
693 may occur during food harvesting. More specific studies on differences in the
694 attachment, stability and potential internalization of viruses should be
695 conducted. These data might be useful when evaluating the risk associated to the
696 consumption of food containing viral genomes.
697
698

5. Conclusions

1. A protocol with a high sensitivity for pathogenic viruses present on the surface of fresh vegetables has been described.
2. The application of viral Metagenomics to water and food safety surveillance is a useful tool for investigating, within a single assay, the potential risk associated to presence of viral pathogens in irrigation river water and vegetable food matrices.
3. New viral sequences related to the *Hepeviridae* and *Picornaviridae* families, which may represent new variants/genera within those families, have been detected. These sequences indicate the merit of further studies.
4. Viral pathogens can represent a threat at low concentrations; thus, high sensitivity to detect low number of viral pathogens is critical for water and food safety. Biases related to the relative viral abundance, the difficulty of obtaining viral concentrates from food and water samples, and the limited availability of viral sequences in public databases are among the problems that remain to be solved. The NGS sequencing output has to be carefully analyzed downstream later on, to improve the taxonomic classification of more robust assemblies.
5. New standardized protocols that will be adjusted to the uniqueness of specific food matrices with the aim of addressing all the aforementioned points are needed to introduce Metagenomics effectively to the water and food safety fields.

Acknowledgments

The study reported here was partially funded by the Programa RecerCaixa 2012 (ACUP-00300) and AGL2011-30461-C02-01/ALI from the Spanish Ministry of Science and Innovation. This study was partially funded by a grant from the Catalan Government to Consolidated Research Group VirBaP (2014SRG914), the JPI Water project METAWATER (4193-00001B) and with the collaboration of the Institut de Recerca de l'Aigua (IdRA). During the development of this study, Xavier Fernandez-Cassi was a fellow of the Catalan Government "AGAUR" (FI-DGR); Natalia Timoneda is a fellow of the Spanish Ministry of Science.

The authors would like to thank the "Servei de Camps Experimentals" at the Universitat de Barcelona, and especially Josep Matas for their assistance with the parsley growth and irrigation. Also we would like to thank Hagar Azab for her English language assistance.

References:

Allen, H.K., Bunge, J., Foster, J.A., Bayles, D.O., Stanton, T.B., 2013. Estimation of viral richness from shotgun metagenomes using a frequency count approach. Microbiome 1, 5. doi:10.1186/2049-2618-1-5

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–10. doi:10.1016/S0022-2836(05)80360-2

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. doi:10.1093/nar/25.17.3389

752 Ambert-Balay, K., Lorrot, M., Bon, F., Giraudon, H., Kaplon, J., Wolfer, M., Lebon,
753 P., Gendrel, D., Pothier, P., 2008. Prevalence and genetic diversity of Aichi
754 virus strains in stool samples from community and hospitalized patients. J.
755 Clin. Microbiol. 46, 1252–1258. doi:10.1128/JCM.02140-07

756 Ao, Y., Yu, J., Li, L., Jin, M., Duan, Z., 2014. Detection of human norovirus GIV.1 in
757 China: a case report. J. Clin. Virol. 61, 298–301.
758 doi:10.1016/j.jcv.2014.08.002

759 Aw, T.G., Wengert, S., Rose, J.B., 2016. Metagenomic analysis of viruses
760 associated with field-grown and retail lettuce identifies human and animal
761 viruses. Int. J. Food Microbiol. 223, 50–56.
762 doi:10.1016/j.ijfoodmicro.2016.02.008

763 Beller, M., Ellis, a, Lee, S.H., Drebot, M. a, Jenkerson, S. a, Funk, E., Sobsey,
764 M.D., Simmons, O.D., Monroe, S.S., Ando, T., Noel, J., Petric, M., Middaugh,
765 J.P., Spika, J.S., 1997. Outbreak of viral gastroenteritis due to a
766 contaminated well. International consequences. JAMA 278, 563–568.
767 doi:10.1001/jama.278.7.563

768 Benson, D.A., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W.,
769 2015. GenBank. Nucleic Acids Res. 43, D30-5. doi:10.1093/nar/gku1216

770 Bernard, H., Faber, M., Wilking, H., Haller, S., Höhle, M., Schielke, A., Ducomble,
771 T., Siffczyk, C., Merbecks, S.S., Fricke, G., Hamouda, O., Stark, K., Werber,
772 D., 2014. Large multistate outbreak of norovirus gastroenteritis associated
773 with frozen strawberries, Germany, 2012. Euro Surveill. 19, 20719.

774 Bofill-Mas, S., Albinana-Gimenez, N., Clemente-Casares, P., Hundesa, A.,
775 Rodriguez-Manzano, J., Allard, A., Calvo, M., Girones, R., 2006.
776 Quantification and stability of human adenoviruses and polyomavirus

777 JCPyV in wastewater matrices. Appl. Environ. Microbiol. 72, 7894–6.
778 doi:10.1128/AEM.00965-06

779 Bofill-Mas, S., Rusiñol, M., Fernandez-Cassi, X., Carratalà, A., Hundesa, A.,
780 Girones, R., 2013. Quantification of human and animal viruses to
781 differentiate the origin of the fecal contamination present in environmental
782 samples. Biomed Res. Int. 2013, 192089. doi:10.1155/2013/192089

783 Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for
784 Illumina sequence data. Bioinformatics 30, 2114–2120.
785 doi:10.1093/bioinformatics/btu170

786 Bosch, A., Pintó, R.M., Guix, S., 2014. Human astroviruses. Clin. Microbiol. Rev.
787 27, 1048–1074. doi:10.1128/CMR.00013-14

788 Brister, J.R., Ako-Adjei, D., Bao, Y., Blinkova, O., 2015. NCBI viral Genomes
789 resource. Nucleic Acids Res. 43, D571–D577. doi:10.1093/nar/gku1207

790 Calgua, B., Fumian, T., Rusiñol, M., Rodriguez-Manzano, J., Mbayed, V.A., Bofill-
791 Mas, S., Miagostovich, M., Girones, R., 2013. Detection and quantification
792 of classic and emerging viruses by skimmed-milk flocculation and PCR in
793 river water from two geographical areas. Water Res. 47, 2797–2810.
794 doi:10.1016/j.watres.2013.02.043

795 Callejón, R.M., Rodríguez-Naranjo, M.I., Ubeda, C., Hornedo-Ortega, R., Garcia-
796 Parrilla, M.C., Troncoso, A.M., 2015. Reported foodborne outbreaks due to
797 fresh produce in the United States and European Union: trends and
798 causes. Foodborne Pathog. Dis. 12, 32–8. doi:10.1089/fpd.2014.1821

799 Cantalupo, P.G., Calgua, B., Zhao, G., 2011. Raw Sewage Harbors Diverse Viral
800 Populations 2, 1–11. doi:10.1128/mBio.00180-11.Editor

801 Coleman, E., Delea, K., Everstine, K., Reimann, D., Ripley, D., 2013. Handling

practices of fresh leafy greens in restaurants: receiving and training. J. Food Prot. 76, 2126–31. doi:10.4315/0362-028X.JFP-13-127

DiCaprio, E., Culbertson, D., Li, J., 2015a. Evidence of the internalization of animal caliciviruses via the roots of growing strawberry plants and dissemination to the fruit. Appl. Environ. Microbiol. 81, 2727–2734. doi:10.1128/AEM.03867-14

DiCaprio, E., Purgianto, A., Ma, Y., Hughes, J., Dai, X., Li, J., 2015b. Attachment and localization of human norovirus and animal caliciviruses in fresh produce. Int. J. Food Microbiol. 211, 101–108. doi:10.1016/j.ijfoodmicro.2015.07.013

Djikeng, A., Kuzmickas, R., Anderson, N.G., Spiro, D.J., 2009. Metagenomic analysis of RNA viruses in a fresh water lake. PLoS One 4. doi:10.1371/journal.pone.0007264

Drexler, J.F., Corman, V.M., Lukashev, A.N., van den Brand, J.M.A., Gmyl, A.P., Brünink, S., Rasche, A., Seggewiß, N., Feng, H., Leijten, L.M., Vallo, P., Kuiken, T., Dotzauer, A., Ulrich, R.G., Lemon, S.M., Drosten, C., 2015. Evolutionary origins of hepatitis A virus in small mammals. Proc. Natl. Acad. Sci. 112, 201516992. doi:10.1073/pnas.1516992112

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, 2526–37. doi:10.1111/j.1462-2920.2012.02791.x

Esseili, M.A., Wang, Q., Saif, L.J., 2012a. Binding of human GII.4 norovirus virus-like particles to carbohydrates of romaine lettuce leaf cell wall materials.

827 Appl. Environ. Microbiol. 78, 786–794. doi:10.1128/AEM.07081-11

828 Esseili, M.A., Wang, Q., Zhang, Z., Saif, L.J., 2012b. Internalization of sapovirus, a
829 surrogate for norovirus, in romaine lettuce and the effect of lettuce latex
830 on virus infectivity. Appl. Environ. Microbiol. 78, 6271–6279.
831 doi:10.1128/AEM.01295-12

832 Ethelberg, S., Lisby, M., B??ttiger, B., Schultz, A.C., Villif, A., Jensen, T., Olsen,
833 K.E., Scheutz, F., Kjels??, C., M??ller, L., 2010. Outbreaks of gastroenteritis
834 linked to lettuce, Denmark, January 2010. Eurosurveillance 15, 1.
835 doi:19484 [pii]

836 Fogeda, M., Avellón, A., Cilla, C.G., Echevarría, J.M., 2009. Imported and
837 autochthonous hepatitis E virus strains in Spain. J. Med. Virol. 81, 1743–
838 1749. doi:10.1002/jmv.21564

839 Gerba, C.P., Goyal, S.M., LaBelle, R.L., Cech, I., Bodgan, G.F., 1979. Failure of
840 indicator bacteria to reflect the occurrence of enteroviruses in marine
841 waters. Am. J. Public Health 69, 1116–9.

842 Greninger, A.L., Naccache, S.N., Federman, S., Yu, G., Mbala, P., Bres, V., Stryke,
843 D., Bouquet, J., Somasekar, S., Linnen, J.M., Dodd, R., Mulembakani, P.,
844 Schneider, B.S., Muyembe-Tamfum, J.-J., Stramer, S.L., Chiu, C.Y., 2015.
845 Rapid metagenomic identification of viral pathogens in clinical samples by
846 real-time nanopore sequencing analysis. Genome Med. 7, 99.
847 doi:10.1186/s13073-015-0220-9

848 Häfliger, D., Hübner, P., Lüthy, J., 2000. Outbreak of viral gastroenteritis due to
849 sewage-contaminated drinking water. Int. J. Food Microbiol. 54, 123–126.
850 doi:10.1016/S0168-1605(99)00176-2

851 Haramoto, E., Katayama, H., Oguma, K., Yamashita, H., Tajima, A., Nakajima, H.,

852 Ohgaki, S., 2006. Seasonal profiles of human noroviruses and indicator
853 bacteria in a wastewater treatment plant in Tokyo, Japan. Water Sci.
854 Technol. 54, 301–8.

855 Hernroth, B.E., Conden-Hansson, A.-C., Rehnstam-Holm, A.-S., Girones, R.,
856 Allard, A.K., 2002. Environmental Factors Influencing Human Viral
857 Pathogens and Their Potential Indicator Organisms in the Blue Mussel,
858 Mytilus edulis: the First Scandinavian Report. Appl. Environ. Microbiol. 68,
859 4523–4533. doi:10.1128/AEM.68.9.4523-4533.2002

860 Hjelmsø, M.H., Hellmér, M., Fernandez-Cassi, X., Timoneda, N., Lukjancenko, O.,
861 Seidel, M., Elsässer, D., Aarestrup, F.M., Löfström, C., Bofill-Mas, S., Abril,
862 J.F., Girones, R., Schultz, A.C., 2017. Evaluation of Methods for the
863 Concentration and Extraction of Viruses from Sewage in the Context of
864 Metagenomic Sequencing. PLoS One 12, e0170199.
865 doi:10.1371/journal.pone.0170199

866 Holm-Hansen, C.C., Midgley, S.E., Fischer, T.K., 2016. Global emergence of
867 enterovirus D68: A systematic review. Lancet Infect. Dis.
868 doi:10.1016/S1473-3099(15)00543-5

869 Holtz, L.R., Finkbeiner, S.R., Zhao, G., Kirkwood, C.D., Girones, R., Pipas, J.M.,
870 Wang, D., 2009. Klassevirus 1, a previously undescribed member of the
871 family Picornaviridae, is globally widespread. Virol. J. 6, 86.
872 doi:10.1186/1743-422X-6-86

873 Hurwitz, B.L., Deng, L., Poulos, B.T., Sullivan, M.B., 2013. Evaluation of methods
874 to concentrate and purify ocean virus communities through comparative,
875 replicated metagenomics. Environ. Microbiol. 15, 1428–40.
876 doi:10.1111/j.1462-2920.2012.02836.x

877 Iritani, N., Yamamoto, S.P., Abe, N., Kubo, H., Oka, T., Kaida, A., 2016. Epidemics
878 of GI.2 sapovirus in gastroenteritis outbreaks during 2012–2013 in Osaka
879 City, Japan. J. Med. Virol. 88, 1187–1193. doi:10.1002/jmv.24451

880 Iwakiri, A., Ganmyo, H., Yamamoto, S., Otao, K., Mikasa, M., Kizoe, S., Katayama,
881 K., Wakita, T., Takeda, N., Oka, T., 2009. Quantitative analysis of fecal
882 sapovirus shedding: Identification of nucleotide substitutions in the capsid
883 protein during prolonged excretion. Arch. Virol. 154, 689–693.
884 doi:10.1007/s00705-009-0358-0

885 Jiang, S.C., 2006. Human Adenoviruses in Water: Occurrence and Health
886 Implications: A Critical Review †. Environ. Sci. Technol. 40, 7132–7140.
887 doi:10.1021/es060892o

888 Karlsson, O.E., Belák, S., Granberg, F., 2013. The effect of preprocessing by
889 sequence-independent, single-primer amplification (SISPA) on
890 metagenomic detection of viruses. Biosecur. Bioterror. 11 Suppl 1, S227-
891 34. doi:10.1089/bsp.2013.0008

892 Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S.,
893 Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B.,
894 Meintjes, P., Drummond, A., 2012. Geneious Basic: An integrated and
895 extendable desktop software platform for the organization and analysis of
896 sequence data. Bioinformatics 28, 1647–1649.
897 doi:10.1093/bioinformatics/bts199

898 Kebschull, J.M., Zador, A.M., 2015. Sources of PCR-induced distortions in high-
899 throughput sequencing data sets. Nucleic Acids Res. 43, e143.
900 doi:10.1093/nar/gkv717

901 Kirk, M.D., Pires, S.M., Black, R.E., Caipo, M., Crump, J.A., Devleesschauwer, B.,

902 Döpfer, D., Fazil, A., Fischer-Walker, C.L., Hald, T., Hall, A.J., Keddy, K.H.,
903 Lake, R.J., Lanata, C.F., Torgerson, P.R., Havelaar, A.H., Angulo, F.J., 2015.
904 World Health Organization Estimates of the Global and Regional Disease
905 Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A
906 Data Synthesis. PLoS Med. 12, e1001921.
907 doi:10.1371/journal.pmed.1001921

908 Kobayashi, S., Fujiwara, N., Yasui, Y., Yamashita, T., Hiramatsu, R., Minagawa, H.,
909 2012. A foodborne outbreak of sapovirus linked to catered box lunches in
910 Japan. Arch. Virol. 157, 1995–7. doi:10.1007/s00705-012-1394-8

911 Kokkinos, P., Kozyra, I., Lazic, S., Bouwknegt, M., Rutjes, S., Willems, K., Moloney,
912 R., de Roda Husman, A.M., Kaupke, A., Legaki, E., D'Agostino, M., Cook, N.,
913 Rzeżutka, A., Petrovic, T., Vantarakis, A., 2012. Harmonised Investigation of
914 the Occurrence of Human Enteric Viruses in the Leafy Green Vegetable
915 Supply Chain in Three European Countries. Food Environ. Virol. 4, 179–
916 191. doi:10.1007/s12560-012-9087-8

917 Koo, H.L., Ajami, N., Atmar, R.L., DuPont, H.L., 2010. Noroviruses: The leading
918 cause of gastroenteritis worldwide. Discov. Med. 10, 61–70.
919 doi:10.1007/s12560-010-9038-1.Noroviruses

920 Kozak, G.K., MacDonald, D., Landry, L., Farber, J.M., 2013. Foodborne outbreaks
921 in Canada linked to produce: 2001 through 2009. J. Food Prot. 76, 173–83.
922 doi:10.4315/0362-028X.JFP-12-126

923 Le Guyader, F.S., Le Saux, J.C., Ambert-Balay, K., Krol, J., Serais, O., Parnaudeau,
924 S., Giraudon, H., Delmas, G., Pommepuy, M., Pothier, P., Atmar, R.L., 2008.
925 Aichi virus, norovirus, astrovirus, enterovirus, and rotavirus involved in
926 clinical cases from a French oyster-related gastroenteritis outbreak. J. Clin.
927 Microbiol. 46, 4011–4017. doi:10.1128/JCM.01044-08

928 Lee, L.E., Cebelinski, E.A., Fuller, C., Keene, W.E., Smith, K., Vinjé, J., Besser, J.M.,
929 2012. Sapovirus outbreaks in long-term care facilities, Oregon and
930 Minnesota, USA, 2002-2009. Emerg. Infect. Dis. 18, 873–876.
931 doi:10.3201/eid1805.111843

932 Legrand-Abravanel, F., Mansuy, J.-M., Dubois, M., Kamar, N., Peron, J.-M.,
933 Rostaing, L., Izopet, J., 2009. Hepatitis E virus genotype 3 diversity, France.
934 Emerg. Infect. Dis. 15, 110–4. doi:10.3201/eid1501.080296

935 Li, T.C., Chijiwa, K., Sera, N., Ishibashi, T., Etoh, Y., Shinohara, Y., Kurata, Y., Ishida,
936 M., Sakamoto, S., Takeda, N., Miyamura, T., 2005. Hepatitis E virus
937 transmission from wild boar meat. Emerg Infect Dis 11, 1958–1960.

938 Marzouk, Y., 1980. Relationship of viruses and indicator bacteria in water and
939 wastewater of Israel. Water Res. 14, 1585–1590. doi:10.1016/0043-
940 1354(80)90062-7

941 Maunula, L., Kaupke, A., Vasickova, P., Söderberg, K., Kozyra, I., Lazic, S., van der
942 Poel, W.H.M., Bouwknegt, M., Rutjes, S., Willems, K.A., Moloney, R.,
943 D'Agostino, M., de Roda Husman, A.M., von Bonsdorff, C.-H., Rzeżutka, A.,
944 Pavlik, I., Petrovic, T., Cook, N., 2013. Tracing enteric viruses in the
945 European berry fruit supply chain. Int. J. Food Microbiol. 167, 177–85.
946 doi:10.1016/j.ijfoodmicro.2013.09.003

947 Namiki, T., Hachiya, T., Tanaka, H., Sakakibara, Y., 2012. MetaVelvet: An
948 extension of Velvet assembler to de novo metagenome assembly from
949 short sequence reads. Nucleic Acids Res. 40. doi:10.1093/nar/gks678

950 Ng, T.F.F., Marine, R., Wang, C., Simmonds, P., Kapusinszky, B., Bodhidatta, L.,
951 Oderinde, B.S., Wommack, K.E., Delwart, E., 2012. High Variety of Known
952 and New RNA and DNA Viruses of Diverse Origins in Untreated Sewage. J.

953 Virol. 86, 12161–12175. doi:10.1128/JVI.00869-12

954 Ng, T.F.F., Sachsenröder, J., Reuter, G., Knowles, N.J., Delwart, E., Johne, R., 2015.
955 Rabovirus: a proposed new picornavirus genus that is phylogenetically
956 basal to enteroviruses and sapeloviruses. Arch. Virol. 160, 2569–2575.
957 doi:10.1007/s00705-015-2523-y

958 Nordgren, J., Matussek, A., Mattsson, A., Svensson, L., Lindgren, P.-E., 2009.
959 Prevalence of norovirus and factors influencing virus concentrations during
960 one year in a full-scale wastewater treatment plant. Water Res. 43, 1117–
961 25. doi:10.1016/j.watres.2008.11.053

962 Oishi, I., Yamazaki, K., Kimoto, T., Minekawa, Y., Utagawa, E., Yamazaki, S.,
963 Inouye, S., Grohmann, G.S., Monroe, S.S., Stine, S.E., Carcamo, C., Ando, T.,
964 Glass, R.I., 1994. A Large Outbreak of Acute Gastroenteritis Associated with
965 Astrovirus among Students and Teachers in Osaka, Japan. J. Infect. Dis. 170,
966 439–443. doi:10.1093/infdis/170.2.439

967 Painter, J.A., Hoekstra, R.M., Ayers, T., Tauxe, R. V, Braden, C.R., Angulo, F.J.,
968 Griffin, P.M., 2013. Attribution of foodborne illnesses, hospitalizations, and
969 deaths to food commodities by using outbreak data, United States, 1998-
970 2008. Emerg. Infect. Dis. 19, 407–15. doi:10.3201/eid1903.111866

971 Park, E.-J., Kim, K.-H., Abell, G.C.J., Kim, M.-S., Roh, S.W., Bae, J.-W., 2011.
972 Metagenomic Analysis of the Viral Communities in Fermented Foods. Appl.
973 Environ. Microbiol. 77, 1284–1291. doi:10.1128/AEM.01859-10

974 Pusch, D., Oh, D.-Y., Wolf, S., Dumke, R., Schröter-Bobsin, U., Höhne, M., Röske,
975 I., Schreier, E., 2005. Detection of enteric viruses and bacterial indicators in
976 German environmental waters. Arch. Virol. 150, 929–47.
977 doi:10.1007/s00705-004-0467-8

978 Räsänen, S., Lappalainen, S., Kaikkonen, S., Hämäläinen, M., Salminen, M.,
979 Vesikari, T., 2010. Mixed viral infections causing acute gastroenteritis in
980 children in a waterborne outbreak. Epidemiol. Infect. 138, 1227–1234.
981 doi:10.1017/S0950268809991671

982 Reuter, G., Boros, Á., Tóth, Z., Gia Phan, T., Delwart, E., Pankovics, P., 2015. A
983 highly divergent picornavirus in an amphibian, the smooth newt
984 (Lissotriton vulgaris). J. Gen. Virol. 96, 2607–13. doi:10.1099/vir.0.000198

985 Riveiro-Barciela, M., Mínguez, B., Gironés, R., Rodriguez-Frías, F., Quer, J., Buti,
986 M., 2015. Phylogenetic Demonstration of Hepatitis E Infection Transmitted
987 by Pork Meat Ingestion. J. Clin. Gastroenterol. 49, 165–168.
988 doi:10.1097/MCG.0000000000000113

989 Rosario, K., Nilsson, C., Lim, Y.W., Ruan, Y., Breitbart, M., 2009. Metagenomic
990 analysis of viruses in reclaimed water. Environ. Microbiol. 11, 2806–2820.
991 doi:10.1111/j.1462-2920.2009.01964.x

992 Sales-Ortells, H., Fernandez-Cassi, X., Timoneda, N., Dürig, W., Girones, R.,
993 Medema, G., 2015. Health risks derived from consumption of lettuces
994 irrigated with tertiary effluent containing norovirus. Food Res. Int. 68, 70–
995 77. doi:10.1016/j.foodres.2014.08.018

996 Sarma, R.H., Saram, M.H. (Mukti H.., State University of New York at Albany.,
997 1990. Structure & methods : proceedings of the Sixth Conversation in
998 the Discipline Biomolecular Stereodynamics held at the State University of
999 New York at Albany, June 6-10, 1989. Adenine Press.

1000 Savichtcheva, O., Okabe, S., 2006. Alternative indicators of fecal pollution:
1001 Relations with pathogens and conventional indicators, current
1002 methodologies for direct pathogen monitoring and future application

perspectives. Water Res. 40, 2463–2476. doi:10.1016/j.watres.2006.04.040

Severi, E., Verhoef, L., Thornton, L., Guzman-Herrador, B.R., Faber, M., Sundqvist, L., Rimhanen-Finne, R., Roque-Afonso, A.M., Ngui, S.L., Allerberger, F., Baumann-Popczyk, A., Muller, L., Parmakova, K., Alfonsi, V., Tavoschi, L., Vennema, H., Fitzgerald, M., Myrmel, M., Gertler, M., Ederth, J., Kontio, M., Vanbockstael, C., Mandal, S., Sadkowska-Todys, M., Tosti, M.E., Schimmer, B., O Gorman, J., Stene-Johansen, K., Wenzel, J.J., Jones, G., Balogun, K., Ciccaglione, A.R., O' Connor, L., Vold, L., Takkinen, J., Rizzo, C., 2015. Large and prolonged food-borne multistate hepatitis A outbreak in Europe associated with consumption of frozen berries, 2013 to 2014. Euro Surveill. Bull. Eur. sur les Mal. Transm. = Eur. Commun. Dis. Bull. 20, 21192.

Smith, D.B., Simmonds, P., Izopet, J., Oliveira-Filho, E.F., Ulrich, R.G., Johne, R., Koenig, M., Jameel, S., Harrison, T.J., Meng, X.J., Okamoto, H., Van Der Poel, W.H.M., Purdy, M.A., 2016. Proposed reference sequences for hepatitis E Virus subtypes. J. Gen. Virol. doi:10.1099/jgv.0.000393

Svraka, S., Vennema, H., Van Der Veer, B., Hedlund, K.O., Thorhagen, M., Siebenga, J., Duizer, E., Koopmans, M., 2010. Epidemiology and genotype analysis of emerging sapovirus-associated infections across Europe. J. Clin. Microbiol. 48, 2191–2198. doi:10.1128/JCM.02427-09

Vega, E., Barclay, L., Gregoricus, N., Williams, K., Lee, D., Vinjé, J., 2011. Novel surveillance network for norovirus gastroenteritis outbreaks, United States. Emerg. Infect. Dis. 17, 1389–1395. doi:10.3201/eid1708.101837

Verhoef, L., Hewitt, J., Barclay, L., Ahmed, S.M., Lake, R., Hall, A.J., Lopman, B., Kroneman, A., Vennema, H., Vinjé, J., Koopmans, M., 2015. Norovirus

genotype profiles associated with foodborne transmission, 1999-2012. Emerg. Infect. Dis. 21, 592–9. doi:10.3201/eid2104.141073

Wang, D., Coscoy, L., Zylberberg, M., Avila, P.C., Boushey, H. a, Ganem, D., DeRisi, J.L., 2002. Microarray-based detection and genotyping of viral pathogens. Proc. Natl. Acad. Sci. U. S. A. 99, 15687–92. doi:10.1073/pnas.242579699

Wang, D., Urisman, A., Liu, Y.T., Springer, M., Ksiazek, T.G., Erdman, D.D., Mardis, E.R., Hickenbotham, M., Magrini, V., Eldred, J., Latreille, J.P., Wilson, R.K., Ganem, D., DeRisi, J.L., 2003. Viral discovery and sequence recovery using DNA microarrays. PLoS Biol. 1. doi:10.1371/journal.pbio.0000002

Wu, F.-T., Oka, T., Takeda, N., Katayama, K., Hansman, G.S., Muo, C.-H., Liang, S.-Y., Hung, C.-H., Dah-Shyong Jiang, D., Hsin Chang, J., Yang, J.-Y., Wu, H.-S., Yang, C.-F., 2008. Acute gastroenteritis caused by GI/2 sapovirus, Taiwan, 2007. Emerg. Infect. Dis. 14, 1169–71. doi:10.3201/eid1407.071531

Yamashita, T., Kobayashi, S., Sakac, K., Nakata, S., Chiba, S., Ishihara, Y., 1991. Isolation of Cytopathic Small Round Viruses with BS-C-l Cells from Patients with Gastroenteritis, The Journal of infectious diseases. doi:10.1093/infdis/164.5.954

Yamashita, Y., Ootsuka, Y., Kondo, R., Oseto, M., Doi, M., Miyamoto, T., Ueda, T., Kondo, H., Tanaka, T., Wakita, T., Katayama, K., Takeda, N., Oka, T., 2010. Molecular characterization of Sapovirus detected in a gastroenteritis outbreak at a wedding hall. J. Med. Virol. 82, 720–6. doi:10.1002/jmv.21646

Yazaki, Y., Mizuo, H., Takahashi, M., Nishizawa, T., Sasaki, N., Gotanda, Y., Okamoto, H., 2003. Sporadic acute or fulminant hepatitis E in Hokkaido,

Japan, may be food-borne, as suggested by the presence of hepatitis E virus in pig liver as food. J. Gen. Virol. 84, 2351–2357. doi:10.1099/vir.0.19242-0

Zerbino, D.R., Birney, E., 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821–829. doi:10.1101/gr.074492.107

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, 303–310. doi:10.1016/j.virol.2014.08.025

Zhang, Y., Zhu, Z., Yang, W., Ren, J., Tan, X., Wang, Y., Mao, N., Xu, S., Zhu, S., Cui, A., Zhang, Y., Yan, D., Li, Q., Dong, X., Zhang, J., Zhao, Y., Wan, J., Feng, Z., Sun, J., Wang, S., Li, D., Xu, W., 2010. An emerging recombinant human enterovirus 71 responsible for the 2008 outbreak of hand foot and mouth disease in Fuyang city of China. Virol. J. 7, 94. doi:10.1186/1743-422X-7-94

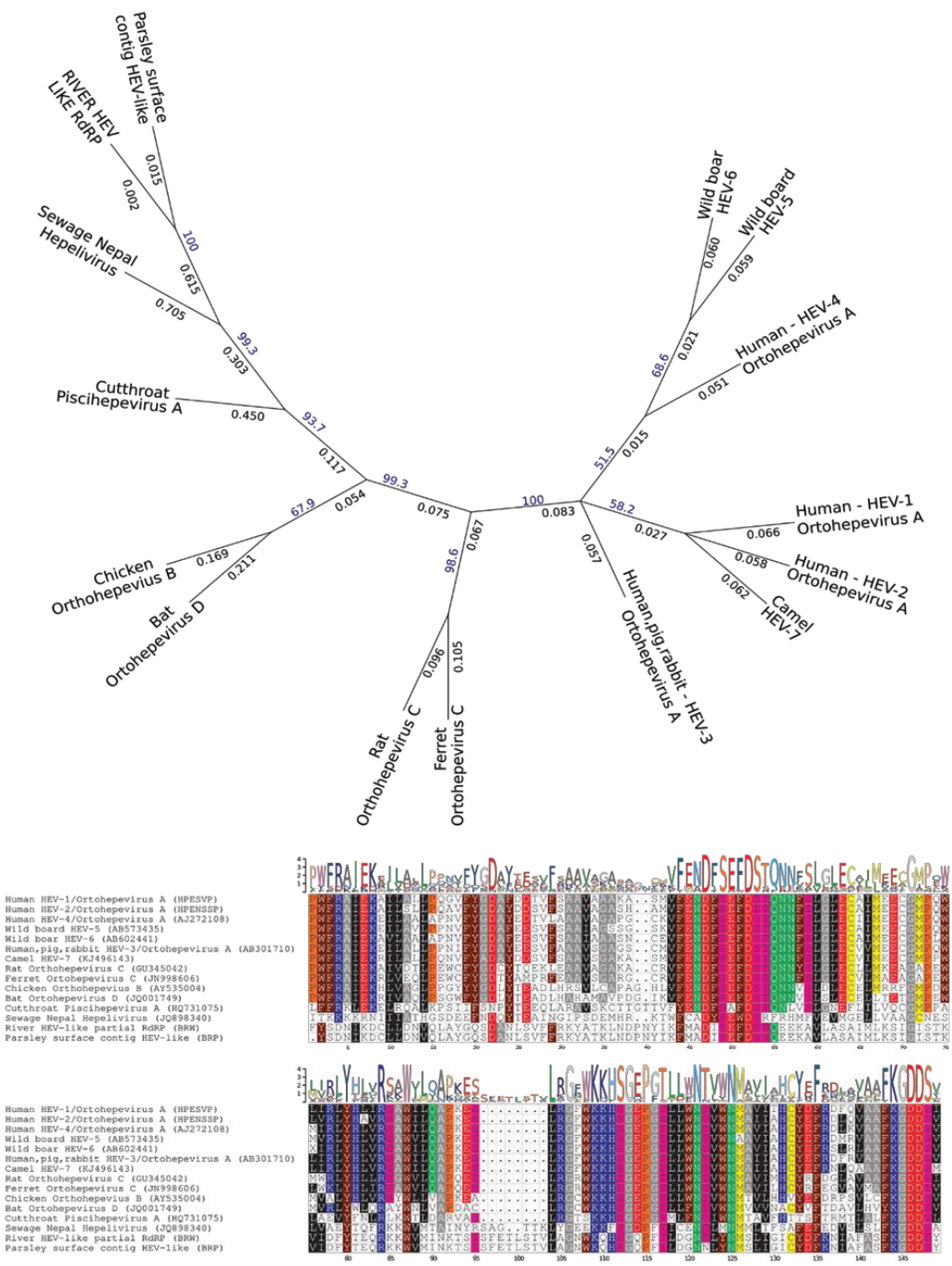


Fig. 1. Phylogenetic analysis of the partial RdRp region for known Hepelivirus and Hepeviridae families, including the River and Parsley surface sequences. A) It is evident that both novel sequence candidates cluster together, and closer to Hepelivirus rather than to Hepeviridae family. Phylogenetic tree was built with Geneious software using Jukes-Cantor model, and clustering method was Neighbor-joining with 1000 bootstrap replicates. Two different scores were included over the branches: bootstrap scores in blue, phylogenetic distances in black. B) The phylogenetic tree was generated from the conserved positions of the RdRp region (204aa), which are shown on this alignment summary produced by Geneious software with default parameters. Sequence accession numbers from GenBank were also annotated with the labels used in the above tree figure. Colors in the alignment correspond to distinct amino acids.

SAMPLES	BRW1		BRW2		BRP		NCP	
Description	Besòs River water		Besòs River water		Parsley Irrigated with river water (20C)		Parsley irrigated with control water	
	Sequences	**Nucleotides**	**Sequences**	**Nucleotides**	**Sequences**	**Nucleotides**	**Sequences**	**Nucleotides**
Raw Reads (MiSeq)	5,426,788	1,450,732,550	2,476,054	577,865,642	1,192,716	327,348,654	8,067,728	2,199,101,314
Clean Reads								
Pair-Ends	3,955,116	930,629,132	2,445,144	508,229,216	329,520	73,820,015	5,301,560	1,264,595,255
Single-Ends	663	123,178	158	26,175	1,808	148,588	866	243,907
Total	3,955,779 72.89%		2,445,302 98.76%		331,328 27.78%		5,302,426 65.72%	
Assembly (MetaVelvet)								
Contigs	243,043	58,899,523	192,193	44,300,670	20,961	5,871,015	10,296	3,160,250
Singletons	2,556,829	599,432,209	1,664,874	339,082,359	191,276	43,401,095	5,169,800	1,233,416,680
$N50_{contigs+singletons}$	1,149,201	270	715,216	232	83,416	324	2,125,927	372
Homology (BLAST)								
Putative viral seqs	77,626 2.77%		44,607 2.40%		59,380 27.98%		3,878,957 74.88%	
Seqs without BLAST hit	2,702,102 96.51%		1,772,659 95.45%		135,759 63.97%		914,807 17.66%	
# Distinct viral families	22		26		25		16	
# Distinct viral species	361		464		139		60	
	Estimated Value	**SE**	**Estimated Value**	**SE**	**Estimated Value**	**SE**	**Estimated Value**	**SE**
Richness	632.4	44.4	923.5	70.9	255.6	34.7	130.3	34.7

Table 1. Metagenomics sequencing summary statistics. Sequences and nucleotide counts are total values; number of pairs is half of the shown values. Percent of reads sequence refer to the total number of raw reads, while the percent of sequences having, or not, a BLAST hit corresponds to the total number of assembled sequences (contigs plus singletons). Difference between the sequences assigned to known viruses and the sequences without a BLAST hit relates to those sequences having a BLAST hit not passing all the filtering criteria for a valid species assignment.

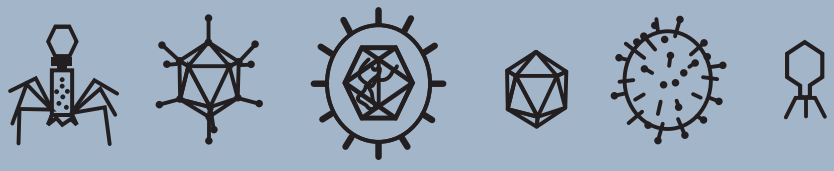
Table 2. BLASTN statistics for human/animal viruses found using metagenomics in Besòs River Water samples.

Sample	Viral family	Viral specie	# related sequences	Maximum contig length	Blast output statistics		
					Larger contig nucleotide identity (%)	Average query coverage (%)	Match GenBank accession number
River water samples used for irrigation (BRW)	*Adenoviridae*	HAdV-41	2	220	100%	100%	KY316164
	Astroviridae	MAstV-1	5	374	94%	99%	KF039911
	Caliciviridae	Goose Calicivirus	2	221	74%	93%	KJ473715
	Picobirnaviridae	Human Picobirnaviridae	2	304	98%	100%	KJ663813
		Porcine Picobirnavirus	2	360	83%	92%	HM070240
	Reoviridae	Human Rotavirus A	6	296	99%	100%	KU048625
	Picornaviridae	Aichi virus	4	385	95%	100%	GQ927712
		Ampivirus	7	716	74%	99%	KP770140
		Bat picornavirus	1	206	77%	87%	HQ595341
		Rat hunnivirus	1	394	91%	100%	KT944214
		Rodent hepatovirus	2	159	80%	99%	KT452641
	Parvoviridae	Kilham rat virus	1	204	79%	80%	AF321230
		Avian adeno-associated virus	5	258	99%	100%	NC_006263
		Caprine Adeno-associated virus	1	376	99%	98%	DQ335246
		Simian adeno-associated virus	1	202	97%	99%	EU285562
		Porcine bocavirus	3	344	87%	100%	KJ622366
		Mouse parvovirus	1	448	84%	100%	KY489986
		Bovine adeno-associated virus	4	274	79%	95%	AY388617
		Rat bocavirus	5	438	94%	100%	KT454517
		Human adeno associated	12	441	98%	100%	AY530578
	Circoviridae	Gull circovirus	3	174	76%	85%	NC_026625
		Avon-Heathcote Estuary associated circular virus 3	2	369	76%	94%	KT454927
		Cyclovirus	13	487	74%	99%	GQ404854
		Human feces pecovirus	2	232	98%	100%	KT600066
		Porcine circovirus	10	148	76%	82%	NC_027796

Table 3. BLASTN statistics for human/vertebrates viruses found using metagenomics in Parsley Plants irrigated with Besòs River water

Sample	Viral family	Viral specie	Number of related sequences	Maximum contig length	Blast output statistics		
					Larger contig nucleotide identity (%)	Average query coverage (%)	Match GenBank accession number
Parsley plants irrigated with river water samples (BRP)	*Caliciviridae*	Human Norovirus GIV.1	4	360	98%	100%	JQ613567
		Human Sapovirus GI.2	16	1780	95%	100%	AB614356
	Hepeviridae	Hepatitis E genotype 3	8	544	86%	99%	JQ953666
	Flaviviridae	Hepatitis G virus	13	1050	93%	100%	KU685422
	Parvoviridae	Human Bocavirus 2a	10	1287	99%	100%	FJ170280
	Anelloviridae	Torque teno midi virus 1	263	306	95%	88%	AB290918
		Torque teno midi virus 2	388	269	77%	87%	AB290919
		Torque teno mini virus 1	44	237	85%	80%	AB026931
		Torque teno mini virus 2	30	241	80%	82%	AB038629
		Torque teno mini virus 3	68	303	88%	89%	AB038630
		Torque teno mini virus 4	2529	315	77%	82%	AB041963
		Torque teno mini virus 5	918	564	91%	83%	AB041962
		Torque teno mini virus 6	280	314	84%	81%	AB026929
		Torque teno mini virus 7	198	297	93%	88%	AB038627
		Torque teno mini virus 8	2503	305	78%	85%	AF291073
		Torque teno mini virus 9	5	189	94%	89%	AB038631
		Torque teno virus 18	1	150	75%	99%	AX025718
		Torque teno virus 2	1	225	75%	95%	AB049608
		Torque teno virus 29	1	386	87%	100%	AB038621
		Torque teno virus 3	2	191	79%	85%	AY666122
	Circoviridae	Avon-Heathcote Estuary associated circular virus 3	1	278	73%	91%	NC_026625
	Picornaviridae	Enterovirus Species A	11	3706	97%	100%	LT617117
		Enterovirus species B	12	394	97%	99%	KU574626

DISCUSSIÓ



DISCUSSIÓ

Els microorganismes transmesos per aigua i aliments són un problema de salut important. Segons estimacions de l'OMS, al món 600 milions de persones, és a dir 1 de cada 10, cau malalta degut al consum d'aliments contaminats produint 420.000 morts cada any (OMS, 2016). La majoria d'aquestes morts es produeixen a països de baixa i mitjana renda on la mala qualitat de l'aigua i la falta d'higiene i sanejament són causa important de morbiditat i mortalitat. S'ha estimat que, en nens, el nombre de morts relacionades amb la mala qualitat de l'aigua és més elevat que el conjunt de defuncions en el mateix col·lectiu causades pel VIH, la malària i el xarampió (Black et al., 2010). En els països amb un bon sistema de sanejament de l'aigua, i on existeixen bones pràctiques agrícoles i de producció d'aliments, es té la percepció de disposar d'aigua i aliments innocus i segurs. En aquests entorns s'han aplicat mesures legislatives que han permès reduir de forma important els problemes microbiològics causats per bacteris, virus i protozous (Newell et al., 2010). Tanmateix, els brots associats al consum d'aigua i aliments contaminats es produeixen també en el nostre entorn de forma recurrent. Entre les causes responsables d'aquests brots podríem apuntar factors com l'alta densitat de població i el seu impacte sobre el medi ambient o la demanda de productes exòtics o fora de temporada que promouen l'existència d'un mercat global. Sovint aquests productes provenen de regions on les condicions de producció i manipulació no compleixen amb els estàndards de qualitat exigibles, ocasionant brots importants (Bernard et al., 2014).

A la regió Mediterrània i en regions del món amb climes més àrids, on la disponibilitat de recursos hídrics és fluctuant i es preveuen períodes llargs de sequera, reduir, reciclar i reutilitzar (RRR) l'aigua disponible és una necessitat. Aquesta idea es veu reforçada per les prediccions relacionades amb el canvi climàtic que preveuen un agreujament d'aquests períodes (IPCC, 2013). La reutilització i reciclatge de les aigües suposa també riscos microbiològics i químics

que han d'ésser pertinentment avaluats (Muñoz et al., 2010). En aquest sentit la caracterització del viroma de l'aigua residual urbana permet determinar els possibles riscos per a la població davant d'eventuals problemes de sobreeiximent, fallada del sistema de clavegueram o la disseminació d'aigua residual a l'ambient insuficientment tractada. Els virus són els principals agents de brots causats per aigua i aliments (Painter et al., 2013).

L'objectiu global plantejat en aquesta tesi ha estat l'anàlisi dels patògens virals que a través de l'aigua residual poden contaminar aigua de reg i aliments.

Per altra banda, la tesi doctoral s'ha centrat en l'avaluació de la qualitat microbiològica de l'aigua regenerada produïda en un sistema de llacunatge mitjançant l'estudi de la reducció de les concentracions de múltiples patògens, posant èmfasi en l'estudi de virus indicadors de contaminació fecal humana i patògens estretament relacionats amb el consum d'aigua i aliments contaminats.

Per altra banda, i amb la finalitat de caracteritzar els riscos virològics associats a l'aigua residual i a la producció d'aigües regenerades s'han desenvolupat tècniques de seqüenciació massiva per l'estudi de virus presents en aquestes matrius. Les aigües regenerades o aigües de riu són una font recurrent per a la irrigació d'aliments, i per tant una possible via d'entrada de patògens a la cadena alimentària.

Avaluació de la qualitat microbiològica de l'aigua regenerada en sistemes de llacunatge

El primer article d'aquesta tesi, és un estudi per avaluar la qualitat microbiològica de l'aigua regenerada produïda a partir d'un sistema de llacunatge. Els mecanismes per produir aigua regenerada es basen en tractaments tecnològicament avançats i normalment costosos que dificulten la seva aplicació en contextos econòmics fràgils o en zones rurals, i especialment en els països més pobres. Per això, alternatives més senzilles, econòmicament més viables i

ambientalment més sostenibles són desitjables. En aquest sentit, els sistemes de llacunatge es consideren una bona alternativa.

El sistema de llacunes estudiat consta de 4 basses connectades, que reben com a aigua d'entrada l'efluent secundari d'una EDAR convencional amb un tractament biològic basat en fangs actius. L'estudi comprèn un any sencer de mostrejos mensuals en què s'han determinat, mitjançant q(RT)PCR i PCR niades, els indicadors virals de contaminació fecal: adenovirus humans (HAdV) i poliomavirus JC (JCPyV), tant a l'entrada com a la sortida del sistema. Per altra banda s'han determinat les concentracions de patògens com norovirus GI (NoV GI) i GII (NoV GII), així com les del virus de l'hepatitis E (VHE). Paral·lelament, s'han avaluat els indicadors bacterians de contaminació fecal (FIB) *E.coli* i enterococs fecals per veure si l'aigua regenerada produïda a la llacuna compleix amb els criteris de qualitat marcats a la normativa (RD 1620/2007). La possibilitat de recreixement d'altres microorganismes amb potencial patogènic, com ara algunes amebes, *Legionella* spp. , *Aeromonas* spp. i *Arcobacter* spp. també ha estat estudiat per altres membres de l'equip.

El sistema de llacunes demostra tenir una capacitat limitada a l'hora d'eliminar els virus presents reduint-ne els valors mitjans de concentració en poc més d'un logaritme. Si només es tenen en consideració els valors obtinguts a partir de tècniques moleculars la capacitat de produir aigua regenerada lliure de virus ha resultat ser força erràtica, presentant resultats similars als descrits per altres autors en sistemes de llacunatge (Jurzik et al., 2015). El nombre de mostres positives per HAdV a l'efluent és més elevat que el nombre de mostres per JCPyV, tot i que els logaritmes de reducció aconseguits per a HAdV són més elevats. Tot i l'elevat nombre de mostres positives a l'efluent, els assajos d'infectivitat per HAdV mitjançant ICC-qPCR van resultar negatius plantejant la possibilitat que no fossin infecciosos. Tot i això s'ha de considerar que els models de cultiu cel·lular no són capaços de detectar tots els virus presents i tenen tendència a infraestimar el nombre de virus (Gerba et al., 2017). Tampoc s'ha detectat cap

positiu per HEV ni a l'entrada del sistema ni a la sortida, probablement per la baixa prevalença d'aquesta infecció en comparació amb altres virus entèrics (Masclaux et al., 2013). Els norovirus humans (HNoV) han presentat els títols més elevats durant els mesos d'hivern, coincidint aquesta època, amb el seu pic estacional. La baixa reducció en còpies genòmiques observada per aquest patògen, freqüentment implicat en brots alimentaris, representa un risc potencial si l'aigua regenerada produïda a la llacuna s'utilitza per al reg de fruites i verdures consumides en cru (0,45 log per NoV GI i 0,72 log per NoV GII). Les reduccions observades pels FIB són superiors a les dels virus arribant a valors propers als 2 logaritmes. Tampoc s'ha detectat una correlació entre els FIB i la presència de virus, evidenciant la poca utilitat que tenen aquests indicadors per a predir la presència de virus humans. Els resultats obtinguts en aquest estudi reforcen la idea que indicadors vírics de contaminació fecal, com HAdV i/o JCPyV, haurien de ser considerats en la legislació com a paràmetres a analitzar en aigües regenerades. Els resultats de reducció de virus observats al sistema de llacunatge són molt similars als obtinguts per altres mètodes de tractament terciari per a produir aigües regenerades i que són econòmicament més costosos com ara la filtració per membrana (≤1 log per NoV (Ottoson et al., 2006), o 1,5 logs per HAdV (Kitajima et al., 2014)), la cloració (0,81-1 log per HAdV (Francy et al., 2012; Kitajima et al., 2014) o la radiació UV (0,24 log per HAdV (Francy et al., 2012)). Tot i la poca legislació referent a virus, existeixen algunes recomanacions per a produir aigües regenerades que suggereixen aconseguir una reducció, d'entre 6-7 $\log_{10}$ al llarg de tot el tractament, si l'aigua regenerada s'ha d'utilitzar com a aigua de reg de productes frescos (Sano et al., 2016; Victor et al., 2008). Assegurant l'assoliment d'aquestes reduccions logarítmiques s'estima que es reduiria el risc en menys d'1:10.000 persones a l'any infectades al consumir aigua o aliments contaminats. Seguint aquestes recomanacions internacionals queda patent que els valors de reducció per virus assolits en el sistema de llacunatge no són suficients per complir les recomanacions, i per tant, l'ús d'efluents d'aquesta

llacuna per a irrigar directament vegetals que es consumeixen crus no seria recomanable. La presència de FIB a l'efluent de la llacuna reafirma aquesta sospita al superar els nivells especificats en el RD1620/2007.

Desenvolupament de protocols per a l'aplicació de tècniques de metagenòmica a l'estudi del viroma en aigua residual

Conèixer el viroma present a l'aigua residual pot oferir informació important des d'una perspectiva ecològica, però també de salut pública, ja que un dels principals *inputs* de l'aigua residual urbana són els virus excretats per la població a partir de les seves femtes, orina o descamació epitelial (Popgeorgiev et al., 2013). L'estudi dels virus presents a l'aigua residual urbana pot oferir informació important sobre les espècies víriques que circulen en una determinada població. Els virus de l'aigua residual poden suposar un problema de salut si l'aigua residual entra en contacte amb aigua i/o aliments, o bé, si els tractaments de les EDAR no aconsegueixen una bona eliminació dels virus de manera que aquests acaben sent disseminats al medi ambient o presents en aigua regenerada.

La utilització de tècniques de seqüenciació massiva per l'estudi del viroma de l'aigua residual ha permès caracteritzar simultàniament, i sota un únic assaig, les diferents famílies víriques presents. Les tècniques de metagenòmica aplicades a virus no es basen en l'ús de cap marcador específic, com el 16S procariòtic o el 18S en eucariotes, sinó que requereixen partir de tot el ADN/ARN present en una determinada mostra (Delwart, 2007). Aquest fet fa que sigui molt important l'eliminació de tot aquell ADN/ARN pròpiament no víric, en especial el ADN bacterià. Sovint la quantitat de ADN/ARN present a les mostres és molt baixa, i es requereixen uns processos d'amplificació independents de seqüència (com el MDA o el SISPA). Com que la metodologia utilitzada és independent de seqüència, aquestes tècniques han permès identificar i caracteritzar noves espècies víriques fins ara desconegudes (Cantalupo et al., 2011; Ng et al., 2012). Aquests factors evidencien que el procés de concentració, extracció i preparació

de les llibreries resulta clau, afectant de forma molt important a la sensibilitat de la tècnica i a la composició del viroma resultant.

En el segon article de la tesi s'avalua l'efecte d'utilitzar diferents mètodes de concentració i extracció d'àcids nucleics per construir les llibreries de seqüenciació, i en el tercer article, es desenvolupa una metodologia més eficient i s'estudia en profunditat el viroma de l'aigua residual.

Una part del primer estudi va realitzar-se durant una estada doctoral a la Universitat Tècnica de Dinamarca (DTU) i inclou l'avaluació d'un total de 16 combinacions diferents de mètodes de concentració de virus i extracció d'àcids nucleics descrits en estudis anteriors amb l'objectiu d'orientar el desenvolupament d'un protocol eficient per l'estudi de metagenòmica de virus a l'aigua. Les mostres van ser dopades amb soques de laboratori de norovirus murí (MNV) i HAdV. Per tal de comprovar la sensibilitat de les tècniques de seqüenciació massiva respecte a tècniques moleculars convencionals, les extraccions van testar-se per q(RT)PCR específiques contra els dos virus mencionats anteriorment. Per a donar robustesa a l'assaig, cadascuna de les combinacions de concentració/extracció va analitzar-se per triplicat. Les 64 llibreries resultants van ser seqüenciades en tres *runs* diferents amb la plataforma illumina Mi-seq de 2x250bp.

Malgrat que els protocols de concentració i extracció estan optimitzats per a virus, un alt percentatge de *reads*, entre el 60-90%, no van mapar amb cap seqüència present a les tres bases de dades de virus utilitzades: genomes complets de virus, seqüències virals extretes de genBank (http://www.ncbi.nlm.nih.gov/genbank/), i la base de dades Vipr (www.ViPRbrc.org) (Pickett et al., 2012). El fet que les bases de dades utilitzades estiguin compostes majoritàriament per seqüències d'origen víric explicaria l'alt percentatge de seqüències que no van poder ser taxonòmicament assignades. Utilitzant un criteri mínim d'identitat del 70%, del total de *reads* que van fer *hit*

contra la base de dades, el 80% va assignar-se taxonòmicament a una espècie vírica.

Un total de 71 famílies víriques diferents van ser detectades a l'aigua residual danesa combinant les 64 llibreries preparades. Seleccionant aquelles famílies víriques que representen més del 99% dels *reads* es va fer una anàlisi de components principals (PCA) per veure l'influencia del procés de concentració i l'extracció en la composició de les poblacions víriques. Els resultats demostren que la influencia del mètode de concentració és més important que la que exerceix el mètode d'extracció tal com evidencia el PCA de la Figura 2 de l'article. Les tres repliques fetes amb el kit nucleospin agrupen de forma conjunta i separades respecte als altres kits d'extracció, independentment del mètode de concentració utilitzat. Aquest agrupament podria explicar-se per l'aplicació d'un pas de DNAsa en columna que hauria eliminat bona part dels virus ADN, fent que el PCA presenti aquest aspecte tan diferencial respecte la resta de kits. El kit de Qiagen presenta una certa variabilitat entre rèpliques. Els kits de Biomerieux i PowerViral sembla que presenten un comportament molt similar, per tant concloem que permeten presentar poblacions properes.

El mètode de concentració amb Polietilenglicol (PEG) inclou un pas de filtració de 0,45 m que eliminaria bona part del ADN/ARN bacterià, incrementant el nombre de *reads* assignats taxonòmicament a virus. De la mateixa manera el percentatge de *reads* virals obtinguts és major quan les mostres s'extreuen amb nucleospin, ja que el pas de DNAsa en columna eliminaria part del ADN bacterià present a les mostres.

L'estimació de la *viral richness* ajuda a entendre i quantificar numèricament l'abundància d'espècies víriques diferents presents en un determinat metagenoma. Aquesta variable es veu més condicionada pel kit d'extracció utilitzat que pel mètode de concentració escollit, sent els kits de Qiagen i PowerViral els que reflecteixen una major diversitat d'espècies virals.

Quan s'avaluen les diferents famílies víriques detectades a l'aigua residual en funció del mètode de concentració i extracció s'observen vàries tendències. El *heatmap* (figura 5) que representa les principals famílies víriques indica que si només interessa estudiar virus ARN els millors resultats s'obtenen a partir de l'ús del kit de nucleospin. Això pot explicar-se per dos motius, per una banda aquest kit conté un pas de DNAsa en columna que hauria eliminat la majoria de ADN bacterià i bona part dels fags que consumirien molts dels *reads*, augmentant la proporció dels virus ARN que infecten humans. Per altra banda els àcids nucleics en aquest kit s'elueixen en un volum menor, fet que afavoriria una major concentració inicial d'àcids nucleics per mL per a preparar la llibreria.

Si el que es pretén és tenir una visió general del viroma present en l'aigua residual, el kit de Qiagen sembla presentar els millors resultats permetent tant l'estudi de virus ADN com ARN. Tenint en compte els mètodes de concentració, sembla que la floculació amb llet descremada (SMF) mostra un major nombre de *reads* associats a patògens transmesos a través d'aigua i aliments contaminats quan es combina amb el kit de nucleospin. Per tant, si haguéssim de recomanar alguna combinació especialment interessant per a l'estudi de virus ARN, principals patògens relacionats amb problemes de salut pública i seguretat alimentaria, la millor combinació seria possiblement SMF amb nucleospin. Malgrat els bons resultats per virus ARN, la SMF aplicada en un estudi anterior ha demostrat tenir un rendiment baix per a famílies ADN, en especial per *Adenoviridae*, *Polyomaviridae* i *Papillomaviridae* on per metagenòmica s'han detectat pocs *reads* associats a aquestes famílies. Especialment rellevant és la baixa detecció dels adenovirus humans, virus amb els que s'havia dopat la mostra, i que han mostrat més baixes quantificacions per qPCR que amb la resta de mètodes de concentració. És possible que el gran volum testat (10 L) hagi provocat una inhibició afectant la quantificació de les mostres i també a la preparació de la llibreria. Karlsson i col·laboradors (2013) també van observar que a major volum de mostra testada, major efecte en la proporció relativa dels

virus, afavorint la detecció per metagenòmica d'aquells virus més abundants. Per altra banda, Bibby i Peccia (2013) van suggerir un possible biaix negatiu de les tècniques de seqüenciació massiva cap a la família *Caliciviridae* que podria explicar la baixa detecció de MNV mitjançant metagenòmica en comparació amb els elevats valors detectats per q(RT)PCR.

Viroma de l'aigua residual urbana

El tercer artícle es titula "estudi del viroma de l'aigua residual de la població utilitzant tècniques de metagenòmica" i ha tingut com a objectius principals definir un protocol eficient per a l'estudi del viroma i caracteritzar els virus humans que el componen. Amb l'objectiu d'estudiar les possibles diferències en la composició del viroma de l'aigua residual urbana segons l'època de l'any, es va recollir una mostra de 10L d'aigua residual als mesos de febrer, maig i setembre. Els virus presents van concentrar-se per SMF i es van quantificar els HAdV per qPCR. Els resultats obtinguts per seqüenciació massiva senyalen una major abundància de virus pertanyents a les famílies *Caliciviridae* i *Astroviridae* durant els mesos d'hivern, coincidint amb els períodes on aquests virus circulen de forma més important per la població (Ahmed et al., 2013; Guix et al., 2002). La majoria d'espècies víriques detectades, independentment de l'estació de l'any, pertanyen a fags de les famílies *Siphoviridae, Myoviridae, Podoviridae* i *Microviridae*, confirmant els resultats d'altres estudis anteriors que assenyalen aquests virus com els organismes més abundants a la terra (Clokie et al., 2011). De manera similar els virus que infecten plantes també representen una part molt important del viroma de l'aigua residual, i en especial, els membres de la família *Virgaviridae*. Aquest fet no hauria de resultar sorprenent si tenim en compte el gran nombre de virus d'origen vegetal que circulen pel tracte digestiu humà i s'excreten en femta (Zhang et al., 2005). Cap adenovirus humà va detectar-se mitjançant metagenòmica en cap de les tres mostres analitzades, tot i els elevats valors obtinguts per qPCR.

Per estudiar amb més profunditat un grup de virus com són els adenovirus es va dissenyar un protocol de metagenòmica dirigida i amb un joc de encebadors per caracteritzar la regió de l'hexó amb prou degeneracions per englobar el nombre més gran d'adenovirus possible. Aquests encebadors van aplicar-se en la mostra que tenia una major concentració d'adenovirus per qPCR. L'amplicó obtingut va ser seqüenciat mitjançant piroseqüenciació. L'ús d'encebadors degenerats va permetre detectar una gran diversitat d'Adenovirus, especialment adenovirus murins i HAdV de l'espècie F, on trobem els HAdV40 i 41, causants de gastroenteritis. L'aplicació de metagenòmica dirigida pot ser un recurs interessant quan es vol estudiar en profunditat la diversitat viral d'una determinada família; ja que actualment les tècniques de metagenòmica no dirigida poden presentar problemes de profunditat o sensibilitat per estudiar en detall la diversitat d'espècies dins d'una mateixa família. Amb l'objectiu d'incrementar la sensibilitat de les tècniques de metagenòmica aplicades es van analitzar de manera comparada dos mètodes de concentració de virus, un d'ells basats en una modificació del protocol de floculació orgànica amb llet descremada analitzant un menor volum de mostra i un protocol basat en la concentració de virus per ultracentrifugació. Les modificacions aplicades al protocol consisteixen en una reducció del volum de mostra inicial a 500 mL, un procés de filtració, i per últim, una reducció en el nombre de cicles de PCR. Per fer una comparació equitativa entre els mètodes es van utilitzar els mateixos mil·lilitres equivalents de mostra d'aigua residual en la preparació de les llibreries.

En conjunt, els dos mètodes van permetre detectar 41 famílies víriques en el total de 4 mostres analitzades detectant la gran majoria de famílies que contenen patògens que es transmeten per aigua i/o aliments contaminats. Les mostres obtingudes per ultracentrifugació van obtenir uns valors de *richness* més elevats que les de floculació. La major part de la diferència observada en els valors de *richness* s'explicaria per la gran quantitat i abundància de bacteriòfags detectats

per ultracentrifugació (Figura 3). Tot i això, determinats virus patògens humans presents en baixes quantitats com els papil·lomavirus i el virus de l'hepatitis E només van ser detectats per ultracentrifugació. El protocol de concentració amb llet descremada, tal com es va observar en l'anterior estudi, va permetre detectar de forma molt eficient virus ARN de les famílies *Caliciviridae, Picornaviridae i Astroviridae* que contenen els principals patògens transmesos pel consum d'aliments i aigües contaminats. Els resultats indiquen que la disminució del volum inicial de mostra, conjuntament amb el procés de filtració i reducció de cicles d'amplificació, ha permès detectar els Adenovirus humans, prèviament no detectats amb el protocol de 10 L. L'efecte de la filtració, la menor presència d'inhibidors atribuïble al menor volum de mostra processat (Schrader et al., 2012) o la reducció en el nombre de cicles de PCR aplicats minimitzant l'amplificació selectiva dels virus més abundants (Karlsson et al., 2013) estarien entre les possibles explicacions. En general, els dos protocols testats han demostrat una eficiència comparable per a l'estudi de virus en aigua residual aplicant tècniques de seqüenciació massiva. El protocol amb menor volum de floculació amb llet descremada ofereix un sistema de concentració de partícules víriques econòmic, ràpid i independent de maquinària complexa demostrant uns resultats equiparables als obtinguts per ultracentrifugació.

En el present estudi també s'ha avaluat la contribució de l'orina en el viroma. Les dades obtingudes, a partir de l'anàlisi de l'orina de 14 individus sans, demostren que el nombre de virus excretats per orina queda restringit majoritàriament a famílies de virus ADN: *Polyomaviridae i Papillomaviridae.* S'han detectat també seqüències de petits virus circulars de cadena senzilla recentment descrits i llunyanament relacionats amb les famílies *Anelloviridae* i *Circoviridae,* que seran objecte d'estudis posteriors.

Estudi metagenòmic de la contaminació viral en vegetals frescos i en aigua de riu utilitzada com a aigua de reg

Els virus representen els principals agents responsables de brots alimentaris. En la cerca d'uns hàbits dietètics més saludables, els consumidors han incrementat el consum de productes frescos, fet que s'ha relacionat amb un augment dels brots per aliments contaminats (Callejón et al., 2015). El nombre de casos de brots associats al consum d'aliments sovint són infravalorats, ja que el seu diagnòstic i notificació és complexa, reportant-se únicament la punta de l'iceberg, en un fenomen conegut com a "piràmide de vigilància"(Nieuwenhuijse and Koopmans, 2017).

En aquest quart article s'ha avaluat el potencial que tenen les tècniques de seqüenciació massiva com a eina de monitorització de la seguretat alimentària. Amb aquesta finalitat, s'han estudiat els virus presents a la superfície de vegetals (juliverts) irrigats amb aigua del riu Besòs, i també presents en la mateixa aigua de reg. Per assegurar que els virus detectats provenien de l'aigua es va regar un grup de plantes control amb una solució nutritiva. Cal destacar que l'aigua de la conca del riu Besòs es troba fortament impactada per l'activitat humana presentant uns nivells elevats de contaminació fecal i que la majoria d'aigua que sustenta el cabal del riu prové dels efluents de les EDAR de la conca.

Després de regar les plantes de julivert durant un mes amb dues mostres diferents d'aigua de riu, i analitzar amb metagenòmica els virus presents, els resultats demostren la presència d'importants patògens humans capaços de produir hepatitis i gastroenteritis en la superfície del vegetal. En total s'han detectat 18 famílies víriques diferents de la superfície del julivert incloent membres de les famílies *Astroviridae, Caliciviridae, Hepeviridae, Picornaviridae* i *Parvoviridae*. D'important rellevància en seguretat alimentària són les seqüències relacionades amb els virus de l'hepatitis E o els NoV GIV. A la mateixa aigua de riu s'han detectat 26 famílies víriques diferents incloent patògens vírics com adenovirus, rotavirus, astrovirus o picornavirus. El sumatori de totes les

seqüències relacionades amb virus patogènics humans no arriba a l'1% dels *reads*. No s'han detectat les mateixes espècies víriques a la superfície dels juliverts que a l'aigua de reg. Això pot explicar-se per dos factors. En primer lloc, l'aplicació de protocols de concentració diferents a causa de la naturalesa diferent de les mostres a processar (superfície vegetal vs. aigua de reg) pot haver influenciat en l'eficiència de recuperació de determinades famílies víriques. En segon lloc, processos d'estabilitat de determinats virus o d'adhesió a la superfície dels vegetals poden haver influït en les espècies virals detectades.

Tot i la importància de les famílies víriques detectades s'ha de tenir present que les tècniques utilitzades només senyalen la presència de genomes vírics que poden no ser infecciosos. Calen més estudis d'estimació de risc que ajudin a comprendre la relació existent entre la presència d'aquests genomes i el risc que impliquen.

L'aplicació de tècniques de seqüenciació massiva al camp de la seguretat alimentària permet, sota una única anàlisi, la detecció de múltiples virus, incloent-hi la detecció d'agents patògens. No obstant això, la presència de patògens en matrius alimentàries sol ser conseqüència d'una contaminació puntual, fent que la seva concentració sigui molt baixa, i per tant, propera als límits de detecció de les tècniques moleculars utilitzades. Per tal d'evitar possibles problemes de sensibilitat, l'ús *d'amplicón sequencing* dels principals patògens implicats en brots alimentaris (HAV, NoV i HEV) conjuntament amb la monitorització de l'aigua de reg per tècniques de metagenòmica no dirigida podrien ser possibles mesures a implementar.

En el segon, tercer i quart article de la tesi s'han requerit estudis bioinformàtics per a la neteja, processament, ensamblat i assignació taxonòmica dels *reads/ contigs* obtinguts. Aquests estudis s'han realitzat amb la col·laboració de Natàlia Timoneda, membre del grup de recerca. Els programes i paràmetres seleccionats durant aquest procés afecten els resultats obtinguts. En el primer estudi de metagenòmica no es van ensamblar els *reads*, sinó que després de filtrar-se en

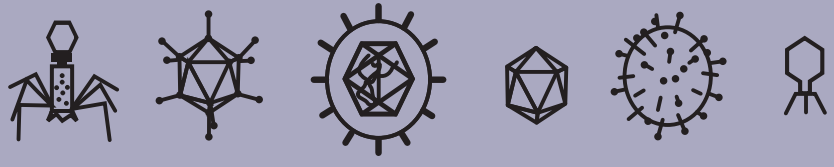
funció de la seva qualitat, aquests van ser mapats directament contra la base de dades. El fet de mapar directament els *reads* disminueix la probabilitat de crear artefactes durant el procés d'ensamblat (Vázquez-Castellanos et al., 2014). Tot ensamblat té com a conseqüència una disminució del nombre de *reads*, i per tant, suposen una pèrdua d'informació. Alguns programes d'ensamblar (com el CLC workbench) donen únicament com a output *contigs* més llargs, sense donar els *reads* no ensamblats. Altres programes (com Meta-Velvet) són més conservadors a l'hora de realitzar aquests processos generant un menor nombre de *contigs* amb una longitud inferior. El procés d'ensamblatge genera seqüències més llargues que ajuden a una correcta anotació taxonòmica de les seqüències i una disminució dels costos computacionals per a fer-ne l'assignació. És important mencionar que l'assignació de seqüències d'un metagenoma depèn molt de la correcta anotació de les seqüències disponibles a les bases de dades. NCBI GenBank conté moltes seqüències mal anotades, un exemple en seria el cas del virus de l'"hepatitis no-A no-B" que en realitat correspon a un bacteriòfag (Cantalupo et al., 2011). L'incorrecta anotació de moltes de les seqüències disponibles explicaria, en part, el baix percentatge de seqüències classificades com a virus presents en els metagenomes, ja que molts bacteriòfags o virus ADN i ARN poden trobar-se integrats en el genoma d'organismes procariotes i/o eucariotes donant una assignació taxonòmica incorrecta. És crucial disposar de genomes ben anotats a les bases de dades de referència per als estudis basats en metagenòmica, i que sovint és necessari contrastar els resultats obtinguts amb més d'una base de dades, per exemple llançar les seqüències contra una base de dades de genomes complets de virus però també contra totes les seqüències víriques presents a GenBank i comparar-ne el resultat.

Els coneixements en bioinformàtica i la gestió i anàlisi de les dades generades per metagenòmica suposa també un problema important per la seva implementació com a tècniques de rutina. Tanmateix, és d'esperar que el desenvolupament de programes cada vegada més intuïtius i visuals, juntament amb l'aparició de noves

tecnologies que permetin la seqüenciació de genomes sencers ajudi a la democratització de l'anàlisi de metagenomes.

L'ús de tècniques de seqüenciació massiva per a l'estudi de virus pot aplicar-se a diferents matrius com aigua residual o aigua de reg utilitzada per al reg d'aliments, constituint una eina interessant per a la realització d'estudis en matèria de seguretat alimentària, d'epidemiologia o vigilància en salut pública.

CONCLUSIONS



CONCLUSIONS

- El sistema de llacunatge estudiat ha demostrat reduir de forma irregular els virus indicadors HAdV, JCPyV i patògens existents en els efluents secundaris de les plantes de tractament d'aigües residuals amb reduccions mitjanes de virus entre 0,45-1,18 log_{10}.

- L'absència de bacteris indicadors fecals a l'efluent de la llacuna no garanteix l'absència de virus tal com s'ha demostrat en altres estudis.

- Els resultats obtinguts en el sistema de llacunatge estudiat indiquen que cal millorar el disseny i la gestió d'aquest tipus de sistema de cara a poder complir amb els requeriments establerts a la normativa i produir una aigua regenerada aplicable al reg de vegetals que es consumeixen crus.

- Els mètodes de concentració i extracció utilitzats poden afectar a la composició viral dels metagenomes, tal com evidencien l'anàlisi de PCR, la viral *richness*, l'especificitat viral o el nombre de virus patògens detectats.

- Variables com el sistema de concentració, el kit d'extracció utilitzat o el volum de mostra processat haurien de ser inclosos en les bases de dades de metagenomes. Aquesta informació pot ser útil per a fer estudis comparatius de meta-anàlisis dels diferents metaviromes.

- S'ha desenvolupat un protocol eficient i sensible per a l'estudi del viroma de l'aigua residual urbana que ha permès detectar importants virus patògens humans de les famílies Par*voviridae, Caliciviridae, Hepeviridae, Adenoviridae, Polyomaviridae, Papillomaviridae* i *Astroviridae*

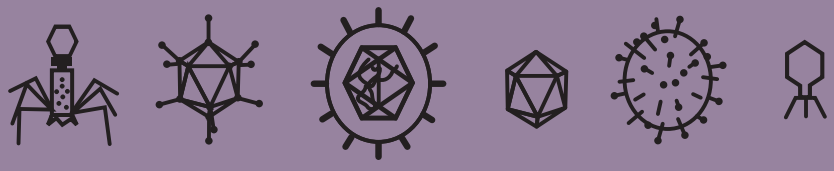
- La utilització de tècniques de seqüenciació massiva a l'aigua residual permet descriure de forma precisa els patògens vírics excretats que circulen per la població i l'ambient constituint una eina important per estudis epidemiològics i de salut pública.

- Els virus detectats en orina són virus de genoma ADN principalment Poliomavirus JC, i en menor quantitat el poliomavirus BK i papil·lomavirus. Seqüències associades a virus circulars ssDNA distantment relacionades amb les famílies *Anelloviridae* i *Circoviridae* han estat detectats i seran considerats per a estudis futurs, ja que podrien constituir nous gèneres/espècies.

- S'ha desenvolupat un protocol sensible per a detectar virus patògens a la superfície de vegetals frescos demostrant el potencial de les tècniques de seqüenciació massiva en el camp de la seguretat alimentària.

- Seqüències distantment relacionades amb les famílies *Hepeviridae* i *Picornaviridae* s'han detectat a l'aigua de reg o a la superfície dels vegetals indicant la presència de possibles nous virus que requeriran ser estudiats en profunditat per a la seva caracterització.

REFERÈNCIES



A

- ACA, Programa de reutilització d'aigua (PRAC), Generalitat de Catalunya, Departament de Medi Ambient i Habitatge, Juny 2009.
- Abd El Galil, K.H., El Sokkary, M.A., Kheira, S.M., Salazar, A.M., Yates, M. V., Chen, W., Mulchandani, A., 2004. Combined Immunomagnetic Separation-Molecular Beacon-Reverse Transcription-PCR Assay for Detection of Hepatitis A Virus from Environmental Samples. Appl. Environ. Microbiol. 70, 4371–4374. doi:10.1128/AEM.70.7.4371-4374.2004
- Adrian, T., Schäfer, G., Cooney, M.K., Fox, J.P., Wigand, R., 1988. Persistent enteral infections with adenovirus types 1 and 2 in infants: no evidence of reinfection. Epidemiol. Infect. 101, 503–9.
- Afiahayati, Sato, K., Sakakibara, Y., 2015. MetaVelvet-SL: an extension of the Velvet assembler to a de novo metagenomic assembler utilizing supervised learning. DNA Res. 22, 69–77. doi:10.1093/dnares/dsu041
- Ahmad, I., Holla, R.P., Jameel, S., 2011. Molecular virology of hepatitis E virus. Virus Res. 161, 47–58. doi:10.1016/j.virusres.2011.02.011
- Ahmed, S.M., Lopman, B.A., Levy, K., 2013. A systematic review and meta-analysis of the global seasonality of norovirus. PLoS One 8, e75922. doi:10.1371/journal.pone.0075922
- Ahmed, W., Stewart, J., Gardner, T., Powell, D., 2008. A real-time polymerase chain reaction assay for quantitative detection of the human-specific enterococci surface protein marker in sewage and environmental waters. Environ. Microbiol. 10, 3255–3264. doi:10.1111/j.1462-2920.2008.01715.x
- Albuquerque, M.C.M., Rocha, L.N., Benati, F.J., Soares, C.C., Maranhão, A.G., Ramírez, M.L., Erdman, D., Santos, N., 2007. Human bocavirus infection in children with gastroenteritis, Brazil. Emerg. Infect. Dis. 13, 1756–8. doi:10.3201/eid1311.070671
- Alfieri, A.A., Alfieri, A.F., Freitas, J.C. de, Silva, C.A. da, Freire, R.L., Barros, A.R.,

Barreiros, M.A.B., Muller, E.E., 1994. Ocurrence of Escherichia Coli, rotavirus, picobirnavirus and Cryptosporidium Parvum in a postweaning diarrhoea focus in swinw. Semin. Ciências Agrárias 15, 5–7. doi:10.5433/1679-0359.1994V15N1P5

- Allander, T., Tammi, M.T., Eriksson, M., Bjerkner, A., Tiveljung-Lindell, A., Andersson, B., 2005. From The Cover: Cloning of a human parvovirus by molecular screening of respiratory tract samples. Proc. Natl. Acad. Sci. 102, 12891–12896. doi:10.1073/pnas.0504666102
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–10. doi:10.1016/S0022-2836(05)80360-2
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. doi:10.1093/nar/25.17.3389
- Ambert-Balay, K., Lorrot, M., Bon, F., Giraudon, H., Kaplon, J., Wolfer, M., Lebon, P., Gendrel, D., Pothier, P., 2008. Prevalence and genetic diversity of Aichi virus strains in stool samples from community and hospitalized patients. J. Clin. Microbiol. 46, 1252–1258. doi:10.1128/JCM.02140-07
- Arana, A., Cilla, G., Montes, M., Gomariz, M., Pérez-Trallero, E., 2014. Genotypes, recombinant forms, and variants of norovirus GII.4 in Gipuzkoa (Basque Country, Spain), 2009-2012. PLoS One 9, e98875. doi:10.1371/journal.pone.0098875
- Arthur, J.L., Higgins, G.D., Davidson, G.P., Givney, R.C., Ratcliff, R.M., 2009. A Novel Bocavirus Associated with Acute Gastroenteritis in Australian Children. PLoS Pathog. 5, e1000391. doi:10.1371/journal.ppat.1000391
- 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. Int. J. Food Microbiol. 223, 50–56. doi:10.1016/j.ijfoodmicro.2016.02.008

B

- Baandrup, L., Thomsen, L.T., Olesen, T.B., Andersen, K.K., Norrild, B., Kjaer, S.K., 2014. The prevalence of human papillomavirus in colorectal adenomas and adenocarcinomas: A systematic review and meta-analysis. Eur. J. Cancer 50, 1446–1461. doi:10.1016/j.ejca.2014.01.019
- Babkin, I. V., Tikunov, A.Y., Zhirakovskaia, E. V., Netesov, S. V., Tikunova, N. V., 2012. High evolutionary rate of human astrovirus. Infect. Genet. Evol. 12, 435–442. doi:10.1016/j.meegid.2012.01.019
- Baert, L., Mattison, K., Loisy-Hamon, F., Harlow, J., Martyres, A., Lebeau, B., Stals, A., Van Coillie, E., Herman, L., Uyttendaele, M., 2011. Review: norovirus prevalence in Belgian, Canadian and French fresh produce: a threat to human health? Int. J. Food Microbiol. 151, 261–9. doi:10.1016/j.ijfoodmicro.2011.09.013
- Baker, K., Morris, J., McCarthy, N., Saldana, L., Lowther, J., Collinson, A., Young, M., 2011. An outbreak of norovirus infection linked to oyster consumption at a UK restaurant, February 2010. J. Public Health (Bangkok). 33, 205–211. doi:10.1093/pubmed/fdq089
- Balkin, H.B., Margolin, A.B., 2010. Detection of poliovirus by ICC/qPCR in concentrated water samples has greater sensitivity and is less costly using BGM cells in suspension as compared to monolayers. Virol. J. 7, 282. doi:10.1186/1743-422X-7-282
- Barrabeig, I., Rovira, A., Buesa, J., Bartolomé, R., Pintó, R., Prellezo, H., Domínguez, À., 2010. Foodborne norovirus outbreak: the role of an asymptomatic food handler. BMC Infect. Dis. 10, 269. doi:10.1186/1471-2334-10-269
- Barril, P.A., Fumian, T.M., Prez, V.E., Gil, P.I., Martínez, L.C., Giordano, M.O., Masachessi, G., Isa, M.B., Ferreyra, L.J., Ré, V.E., Miagostovich, M., Pavan, J.V., Nates, S.V., 2015. Rotavirus seasonality in urban sewage from Argentina: Effect of meteorological variables on the viral load and the genetic diversity. Environ. Res. 138, 409–415. doi:10.1016/j.envres.2015.03.004

- Barouch, D.H., Kik, S. V., Weverling, G.J., Dilan, R., King, S.L., Maxfield, L.F., Clark, S., Ng'ang'a, D., Brandariz, K.L., Abbink, P., Sinangil, F., de Bruyn, G., Gray, G.E., Roux, S., Bekker, L.-G., Dilraj, A., Kibuuka, H., Robb, M.L., Michael, N.L., Anzala, O., Amornkul, P.N., Gilmour, J., Hural, J., Buchbinder, S.P., Seaman, M.S., Dolin, R., Baden, L.R., Carville, A., Mansfield, K.G., Pau, M.G., Goudsmit, J., 2011. International seroepidemiology of adenovirus serotypes 5, 26, 35, and 48 in pediatric and adult populations. Vaccine 29, 5203–5209. doi:10.1016/j.vaccine.2011.05.025
- Batts, W., Yun, S., Hedrick, R., Winton, J., 2011. A novel member of the family Hepeviridae from cutthroat trout (Oncorhynchus clarkii). Virus Res. 158, 116–123. doi:10.1016/j.virusres.2011.03.019
- Baylis, S.A., Tuke, P.W., Miyagawa, E., Blümel, J., 2013. Studies on the inactivation of human parvovirus 4. Transfusion 53, 2585–2592. doi:10.1111/trf.12372
- Belfort, G., Paluszek, A., Sturman, L.S., 1982. Enterovirus Concentration Using Automated Hollow Fiber Ultrafiltration. Water Sci. Technol. 14.
- Beller, M., Ellis, a, Lee, S.H., Drebot, M. a, Jenkerson, S. a, Funk, E., Sobsey, M.D., Simmons, O.D., Monroe, S.S., Ando, T., Noel, J., Petric, M., Middaugh, J.P., Spika, J.S., 1997. Outbreak of viral gastroenteritis due to a contaminated well. International consequences. JAMA 278, 563–568. doi:10.1001/jama.278.7.563
- Bernard, H.-U., Burk, R.D., Chen, Z., van Doorslaer, K., zur Hausen, H., de Villiers, E.-M., 2010. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. Virology 401, 70–9. doi:10.1016/j.virol.2010.02.002
- Bernard, H., Faber, M., Wilking, H., Haller, S., Höhle, M., Schielke, A., Ducomble, T., Siffczyk, C., Merbecks, S.S., Fricke, G., Hamouda, O., Stark, K., Werber, D., 2014. Large multistate outbreak of norovirus gastroenteritis associated with frozen strawberries, Germany, 2012. Euro Surveill. 19, 20719.
- Bhattacharya, R., Sahoo, G.C., Nayak, M.K., Rajendran, K., Dutta, P., Mitra, U., Bhattacharya, M.K., Naik, T.N., Bhattacharya, S.K., Krishnan, T., 2007. Detection of

Genogroup I and II human picobirnaviruses showing small genomic RNA profile causing acute watery diarrhoea among children in Kolkata, India☆. Infect. Genet. Evol. 7, 229–238. doi:10.1016/j.meegid.2006.09.005

- **Bibby, K., Peccia, J.,** 2013a. Identification of viral pathogen diversity in sewage sludge by metagenome analysis. Environ. Sci. Technol. 47, 1945–51. doi:10.1021/es305181x
- **Bibby, K., Viau, E., Peccia, J.,** 2011. Viral metagenome analysis to guide human pathogen monitoring in environmental samples. Lett. Appl. Microbiol. 52, 386–92. doi:10.1111/j.1472-765X.2011.03014.x
- **Bishop, R.F.,** 1996. Natural history of human rotavirus infection. Arch. Virol. Suppl. 12, 119–28.
- **Bisson, J.W., Cabelli, V.J.,** 1980. Clostridium perfringens as a water pollution indicator. J. Water Pollut. Control Fed. 52, 241–8.
- **Black, R.E., Cousens, S., Johnson, H.L., Lawn, J.E., Rudan, I., Bassani, D.G., Jha, P., Campbell, H., Walker, C.F., Cibulskis, R., Eisele, T., Liu, L., Mathers, C., Child Health Epidemiology Reference Group of WHO and UNICEF,** 2010. Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet 375, 1969–1987. doi:10.1016/S0140-6736(10)60549-1
- **Blinkova, O., Rosario, K., Li, L., Kapoor, A., Slikas, B., Bernardin, F., Breitbart, M., Delwart, E.,** 2009. Frequent detection of highly diverse variants of cardiovirus, cosavirus, bocavirus, and circovirus in sewage samples collected in the United States. J. Clin. Microbiol. 47, 3507–13. doi:10.1128/JCM.01062-09
- **Block, J.C., Schwartzbrod, L.,** 1989. Viruses in water systems : detection and identification. VCH, New York N.Y. ;Weinheim Federal Republic of Germany.
- **Boccia, D., Guthmann, J.-P., Klovstad, H., Hamid, N., Tatay, M., Ciglenecki, I., Nizou, J.-Y., Nicand, E., Guerin, P.J.,** 2006. High mortality associated with an outbreak of hepatitis E among displaced persons in Darfur, Sudan. Clin. Infect. Dis. 42, 1679–84. doi:10.1086/504322
- **Bodewes, R., Ruiz-Gonzalez, A., Schapendonk, C.M., van den Brand, J.M.,**

Osterhaus, A.D., Smits, S.L., 2014a. Viral metagenomic analysis of feces of wild small carnivores. Virol. J. 11, 89. doi:10.1186/1743-422X-11-89

- Bodewes, R., Ruiz-Gonzalez, A., Schürch, A.C., Osterhaus, A.D.M.E., Smits, S.L., 2014b. Novel Divergent Rhabdovirus in Feces of Red Fox, Spain. Emerg. Infect. Dis. 20, 2172–2174. doi:10.3201/eid2012.140236
- 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. 72, 7894–6. doi:10.1128/AEM.00965-06
- Bofill-Mas, S., Calgua, B., Clemente-Casares, P., La Rosa, G., Iaconelli, M., Muscillo, M., Rutjes, S., de Roda Husman, A.M., Grunert, A., Gräber, I., Verani, M., Carducci, A., Calvo, M., Wyn-Jones, P., Girones, R., 2010. Quantification of Human Adenoviruses in European Recreational Waters. Food Environ. Virol. 2, 101–109. doi:10.1007/s12560-010-9035-4
- Bofill-Mas, S., Pina, S., Girones, R., 2000. Documenting the epidemiologic patterns of polyomaviruses in human populations by studying their presence in urban sewage. Appl. Environ. Microbiol. 66, 238–45.
- Boisvert, S., Raymond, F., Godzaridis, E., Laviolette, F., Corbeil, J., 2012. Ray Meta: scalable de novo metagenome assembly and profiling. Genome Biol. 13, R122. doi:10.1186/gb-2012-13-12-r122
- Bonilla, N., Santiago, T., Marcos, P., Urdaneta, M., Domingo, J.S., Toranzos, G.A., 2010. Enterophages, a group of phages infecting Enterococcus faecalis, and their potential as alternate indicators of human faecal contamination. Water Sci. Technol. 61, 293. doi:10.2166/wst.2010.815
- Boros, ??kos, Kiss, T., Kiss, O., Pankovics, P., Kapusinszky, B., Delwart, E., Reuter, G., 2013. Genetic characterization of a novel picornavirus distantly related to the marine mammal-infecting aquamaviruses in a long-distance migrant bird species, European roller (Coracias garrulus). J. Gen. Virol. 94, 2029–2035. doi:10.1099/vir.0.054676-0

- **Bosch, A., Pintó, R.M., Guix, S.,** 2014a. Human Astroviruses. Clin. Microbiol. Rev. 27, 1048–1074. doi:10.1128/CMR.00013-14
- **Bosch, A., Pintó, R.M., Villena, C., Abad, F.X.**, 1997. Persistence of human astrovirus in fresh and marine water, in: Water Science and Technology. pp. 243–247. doi:10.1016/S0273-1223(97)00266-7
- **Boxman, I., Dijkman, R., Verhoef, L., Maat, A., van Dijk, G., Vennema, H., Koopmans, M.,** 2009. Norovirus on swSARVIKIVI, E., ROIVAINEN, M., MAUNULA, L., NISKANEN, T., KORHONEN, T., LAPPALAINEN, M., KUUSI, M., 2012. Multiple norovirus outbreaks linked to imported frozen raspberries. Epidemiol. Infect. 140, 260–267. doi:10.1017/S0950268811000379abs t. J. Food Prot. 72, 1753–5.
- **Browning, G.F., Chalmers, R.M., Snodgrass, D.R., Batt, R.M., Hart, C.A., Ormarod, S.E., Leadon, D., Stoneham, S.J., Rossdale, P.D.,** 1991. The prevalence of enteric pathogens in diarrhoeic thoroughbred foals in Britain and Ireland. Equine Vet. J. 23, 405–9.
- **Buck, C.B., Phan, G.Q., Raiji, M.T., Murphy, P.M., McDermott, D.H., McBride, A.A.,** 2012. Complete genome sequence of a tenth human polyomavirus. J. Virol. 86, 10887. doi:10.1128/JVI.01690-12
- **Bull, R.A., Tanaka, M.M., White, P.A.,** 2007. Norovirus recombination. J. Gen. Virol. 88, 3347–3359. doi:10.1099/vir.0.83321-0

C

- **Calgua, B., Fumian, T., Rusiñol, M., Rodriguez-Manzano, J., Mbayed, V.A., Bofill-Mas, S., Miagostovich, M., Girones, R.,** 2013a. Detection and quantification of classic and emerging viruses by skimmed-milk flocculation and PCR in river water from two geographical areas. Water Res. 47, 2797–2810. doi:10.1016/j.watres.2013.02.043
- **Calgua, B., Mengewein, A., Grunert, A., Bofill-Mas, S., Clemente-Casares, P., Hundesa, A., Wyn-Jones, A.P., López-Pila, J.M., Girones, R.,** 2008. Development and application of a one-step low cost procedure to concentrate viruses from

seawater samples. J. Virol. Methods 153, 79–83. doi:10.1016/j.jviromet.2008.08.003

- Calgua, B., Rodriguez-Manzano, J., Hundesa, A., Suñen, E., Calvo, M., Bofill-Mas, S., Girones, R., 2013c. New methods for the concentration of viruses from urban sewage using quantitative PCR. J. Virol. Methods 187, 215–21. doi:10.1016/j.jviromet.2012.10.012
- 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 Pathog. Dis. 12, 32–8. doi:10.1089/fpd.2014.1821
- Calvignac-Spencer, S., Feltkamp, M.C.W., Daugherty, M.D., Moens, U., Ramqvist, T., Johne, R., Ehlers, B., Ehlers, B., 2016. A taxonomy update for the family Polyomaviridae. Arch. Virol. 161, 1739–1750. doi:10.1007/s00705-016-2794-y
- Cannon, R.O., Poliner, J.R., Hirschhorn, R.B., Rodeheaver, D.C., Silverman, P.R., Brown, E.A., Talbot, G.H., Stine, S.E., Monroe, S.S., Dennis, D.T., 1991. A multistate outbreak of Norwalk virus gastroenteritis associated with consumption of commercial ice. J. Infect. Dis. 164, 860–3.
- Cantalupo, P.G., Calgua, B., Zhao, G., 2011. Raw Sewage Harbors Diverse Viral Populations 2, 1–11. doi:10.1128/mBio.00180-11.Editor
- Carducci, A., Caponi, E., Ciurli, A., Verani, M., 2015. Possible Internalization of an Enterovirus in Hydroponically Grown Lettuce. Int. J. Environ. Res. Public Health 12, 8214–27. doi:10.3390/ijerph120708214
- Carratalà, A., Rodriguez-Manzano, J., Hundesa, A., Rusiñol, M., Fresno, S., Cook, N., Girones, R., 2013. Effect of temperature and sunlight on the stability of human adenoviruses and MS2 as fecal contaminants on fresh produce surfaces. Int. J. Food Microbiol. 164, 128–34. doi:10.1016/j.ijfoodmicro.2013.04.007
- Carter, M.J., 2005. Enterically infecting viruses: pathogenicity, transmission and significance for food and waterborne infection. J. Appl. Microbiol. 98, 1354–80. doi:10.1111/j.1365-2672.2005.02635.x

- Casas-Alba, D., de Sevilla, M.F., Valero-Rello, A., Fortuny, C., García-García, J.-J., Ortez, C., Muchart, J., Armangué, T., Jordan, I., Luaces, C., Barrabeig, I., González-Sanz, R., Cabrerizo, M., Muñoz-Almagro, C., Launes, C., 2017. Outbreak of brainstem encephalitis associated with enterovirus-A71 in Catalonia, Spain (2016): a clinical observational study in a children's reference centre in Catalonia. Clin. Microbiol. Infect. 142, 776–788. doi:10.1016/j.cmi.2017.03.016
- Cascio, A., Bosco, M., Vizzi, E., Giammanco, A., Ferraro, D., Arista, S., 1996. Identification of picobirnavirus from faeces of Italian children suffering from acute diarrhea. Eur. J. Epidemiol. 12, 545–547.
- CDC, 2000. Foodborne outbreak of Group A rotavirus gastroenteritis among college students--District of Columbia, March-April 2000. MMWR Morb. Mortal. Wkly. Rep. 49, 1131–1133.
- Chen, H., Chen, X.-Z., Waterboer, T., Castro, F.A., Brenner, H., 2015. Viral infections and colorectal cancer: A systematic review of epidemiological studies. Int. J. Cancer 137, 12–24. doi:10.1002/ijc.29180
- Chen, H., Yao, Y., Liu, X., Xiao, N., Xiao, Y., Huang, Y., Chen, Q., Yu, S., 2014. Molecular detection of human parechovirus in children with acute gastroenteritis in Guangzhou, China. Arch. Virol. 159, 971–977. doi:10.1007/s00705-013-1915-0
- Chirmule, N., Propert, K.J., Magosin, S.A., Qian, Y., Qian, R., Wilson, J.M., 1999. Immune responses to adenovirus and adeno-associated virus in humans. Gene Ther. 6, 1574–1583. doi:10.1038/sj.gt.3300994
- Clokie, M.R., Millard, A.D., Letarov, A. V, Heaphy, S., 2011. Phages in nature. Bacteriophage 1, 31–45. doi:10.4161/bact.1.1.14942
- Collier, M.G., Khudyakov, Y.E., Selvage, D., Adams-Cameron, M., Epson, E., Cronquist, A., Jervis, R.H., Lamba, K., Kimura, A.C., Sowadsky, R., Hassan, R., Park, S.Y., Garza, E., Elliott, A.J., Rotstein, D.S., Beal, J., Kuntz, T., Lance, S.E., Dreisch, R., Wise, M.E., Nelson, N.P., Suryaprasad, A., Drobeniuc, J., Holmberg, S.D., Xu, F., Hepatitis A Outbreak Investigation Team, 2014. Outbreak of hepatitis A in the USA associated with frozen pomegranate arils imported from Turkey: an

epidemiological case study. Lancet Infect. Dis. 14, 976–981. doi:10.1016/S1473-3099(14)70883-7

- Cotmore, S.F., Agbandje-McKenna, M., Chiorini, J.A., Mukha, D. V, Pintel, D.J., Qiu, J., Soderlund-Venermo, M., Tattersall, P., Tijssen, P., Gatherer, D., Davison, A.J., 2014. The family Parvoviridae. Arch. Virol. 159, 1239–47. doi:10.1007/s00705-013-1914-1

D

- D'Ambrosio, E., Del Grosso, N., Chicca, A., Midulla, M., 1982. Neutralizing antibodies against 33 human adenoviruses in normal children in Rome. J. Hyg. (Lond). 89, 155–61.
- Da Silva, A.K., Le Saux, J.C., Parnaudeau, S., Pommepuy, M., Elimelech, M., Le Guyader, F.S., 2007. Evaluation of removal of noroviruses during wastewater treatment, using real-time reverse transcription-PCR: Different behaviors of genogroups I and II. Appl. Environ. Microbiol. 73, 7891–7897. doi:10.1128/AEM.01428-07
- Damin, D.C., Ziegelmann, P.K., Damin, A.P., 2013. Human papillomavirus infection and colorectal cancer risk: a meta-analysis. Color. Dis. 15, e420–e428. doi:10.1111/codi.12257
- Dare, A., 2015. Irrigation with treated municipal wastewater in Indiana, United States. J. Soil Water Conserv. 70, 89A–94A. doi:10.2489/jswc.70.4.89A
- De Benedictis, P., Schultz-Cherry, S., Burnham, A., Cattoli, G., 2011. Astrovirus infections in humans and animals – Molecular biology, genetic diversity, and interspecies transmissions. Infect. Genet. Evol. 11, 1529–1544. doi:10.1016/j.meegid.2011.07.024
- de Graaf, M., van Beek, J., Vennema, H., Podkolzin, A.T., Hewitt, J., Bucardo, F., Templeton, K., Mans, J., Nordgren, J., Reuter, G., Lynch, M., Rasmussen, L.D., Iritani, N., Chan, M.C., Martella, V., Ambert-Balay, K., Vinjé, J., White, P.A., Koopmans, M.P., 2015. Emergence of a novel GII.17 norovirus – End of the GII.4 era? Euro

Surveill. 20.

- Delwart, E.L., 2007. Viral metagenomics. Rev. Med. Virol. 17, 115–131. doi:10.1002/rmv.532
- Dennehy, P.H., 2000. Transmission of rotavirus and other enteric pathogens in the home. Pediatr. Infect. Dis. J. 19, S103-5.
- Di Bonito, P., Della Libera, S., Petricca, 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. doi:10.1099/vir.0.071787-0
- DiCaprio, E., Purgianto, A., Ma, Y., Hughes, J., Dai, X., Li, J., 2015. Attachment and localization of human norovirus and animal caliciviruses in fresh produce. Int. J. Food Microbiol. 211, 101–108. doi:10.1016/j.ijfoodmicro.2015.07.013
- Directiva 98/83/CE, 1998. Calidad de las aguas destinadas al consumo humano
- Directiva 91/271/CE, 1991.Tratamiento de las aguas residuales urbanas
- Directiva 2006/7/CE, 2006. Gestión de la calidad de las aguas de baño
- Divizia, M., Santi, A.L., Panà, A., 1989. Ultrafiltration: an efficient second step for hepatitis A virus and poliovirus concentration. J. Virol. Methods 23, 55–62. doi:10.1016/0166-0934(89)90089-X
- Djikeng, A., Kuzmickas, R., Anderson, N.G., Spiro, D.J., 2009. Metagenomic analysis of RNA viruses in a fresh water lake. PLoS One 4. doi:10.1371/journal.pone.0007264
- Doceul, V., Bagdassarian, E., Demange, A., Pavio, N., 2016. Zoonotic Hepatitis E Virus: Classification, Animal Reservoirs and Transmission Routes. Viruses 8, 270. doi:10.3390/v8100270
- Dong, Y., Kim, J., Lewis, G.D., 2010. Evaluation of methodology for detection of human adenoviruses in wastewater, drinking water, stream water and recreational waters. J. Appl. Microbiol. 108, 800–809. doi:10.1111/j.1365-2672.2009.04477.x
- Doorbar, J., Quint, W., Banks, L., Bravo, I.G., Stoler, M., Broker, T.R., Stanley, M.A.,

2012. The biology and life-cycle of human papillomaviruses. Vaccine 30 Suppl 5, F55-70. doi:10.1016/j.vaccine.2012.06.083

- Dowell, S.F., Groves, C., Kirkland, K.B., Cicirello, H.G., Ando, T., Jin, Q., Gentsch, J.R., Monroe, S.S., Humphrey, C.D., Slemp, C., 1995. A multistate outbreak of oyster-associated gastroenteritis: implications for interstate tracing of contaminated shellfish. J. Infect. Dis. 171, 1497–503.

E

- EBI Metagenomics website. (Accesed 12.01.107). https://www.ebi.ac.uk/metagenomics/;jsessionid=83AC224E486AE89FB7FE8CF5F8548A56
- Edwards, R.A., Rohwer, F., 2005. Opinion: Viral metagenomics. Nat. Rev. Microbiol. 3, 504–510. doi:10.1038/nrmicro1163
- Enriquez, C., 1995. Survival of the enteric adenoviruses 40 and 41 in tap, sea, and waste water. Water Res. 29, 2548–2553. doi:10.1016/0043-1354(95)00070-2
- Estes, M.K., Graham, D.Y., Smith, E.M., Gerba, C.P., 1979. Rotavirus Stability and Inactivation. J. Gen. Virol. 43, 403–409. doi:10.1099/0022-1317-43-2-403
- Ethelberg, S., Lisby, M., Bottiger, B., Schultz, A.C., Villif, A., Jensen, T., Olsen, K.E., Scheutz, F., Kjelso, C., Muller, L., 2010. Outbreaks of gastroenteritis linked to lettuce, Denmark, January 2010. Euro Surveill. 15.
- Ettayebi, K., Crawford, S.E., Murakami, K., Broughman, J.R., Karandikar, U., Tenge, V.R., Neill, F.H., Blutt, S.E., Zeng, X.-L., Qu, L., Kou, B., Opekun, A.R., Burrin, D., Graham, D.Y., Ramani, S., Atmar, R.L., Estes, M.K., 2016. Replication of human noroviruses in stem cell–derived human enteroids. Science (80-.).
- Ewing, B., Hillier, L., Wendl, M.C., Green, P., 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 8, 175–85.

F

- **Family – Astroviridae,** 2012. , in: Virus Taxonomy. pp. 953–959. doi:10.1016/B978-0-12-384684-6.00081-1
- **Family – Circoviridae,** 2012. , in: Virus Taxonomy. pp. 343–349. doi:10.1016/B978-0-12-384684-6.00034-3
- **Farkas, T., Sestak, K., Wei, C., Jiang, X.**, 2008. Characterization of a Rhesus Monkey Calicivirus Representing a New Genus of Caliciviridae. J. Virol. 82, 5408–5416. doi:10.1128/JVI.00070-08
- **Farrah, S.R., Gerba, C.P., Wallis, C., Melnick, J.L.,** 1976. Concentration of viruses from large volumes of tap water using pleated membrane filters. Appl. Environ. Microbiol. 31, 221–6.
- **Farrah, S.R., Preston, D.R.,** 1985. Concentration of viruses from water by using cellulose filters modified by in situ precipitation of ferric and aluminum hydroxides. Appl. Environ. Microbiol. 50, 1502–4.
- **Felsenstein, S., Yang, S., Eubanks, N., Sobrera, E., Grimm, J.P., Aldrovandi, G.,** 2014. Human Parechovirus Central Nervous System Infections in Southern California Children. Pediatr. Infect. Dis. J. 33, e87–e91. doi:10.1097/INF.0000000000000112
- **Finkbeiner, S.R., Holtz, L.R., Jiang, Y., Rajendran, P., Franz, C.J., Zhao, G., Kang, G., Wang, D.,** 2009a. Human stool contains a previously unrecognized diversity of novel astroviruses. Virol. J. 6, 161. doi:10.1186/1743-422X-6-161
- **Finkbeiner, S.R., Kirkwood, C.D., Wang, D.,** 2008. Complete genome sequence of a highly divergent astrovirus isolated from a child with acute diarrhea. Virol. J. 5, 117. doi:10.1186/1743-422X-5-117
- **Finkbeiner, S.R., Li, Y., Ruone, S., Conrardy, C., Gregoricus, N., Toney, D., Virgin, H.W., Anderson, L.J., Vinjé, J., Wang, D., Tong, S.,** 2009b. Identification of a novel astrovirus (astrovirus VA1) associated with an outbreak of acute gastroenteritis. J. Virol. 83, 10836–9. doi:10.1128/JVI.00998-09

- Fischer, T., Nielsen, A., Sydenham, T., Andersen, P., Andersen, B., Midgley, S., 2014. Emergence of enterovirus 71 C4a in Denmark, 2009 to 2013. Eurosurveillance 19, 20911. doi:10.2807/1560-7917.ES2014.19.38.20911
- Fleischmann, R., Adams, M., White, O., Clayton, R., Kirkness, E., Kerlavage, A., Bult, C., Tomb, J., Dougherty, B., Merrick, J., al., et, 1995. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science (80-.). 269.
- Fleisher, J.M., Kay, D., Salmon, R.L., Jones, F., Wyer, M.D., Godfree, A.F., 1996. Marine waters contaminated with domestic sewage: nonenteric illnesses associated with bather exposure in the United Kingdom. Am. J. Public Health 86, 1228–34.
- Flotte, T.R., Berns, K.I., 2005. Adeno-Associated Virus: A Ubiquitous Commensal of Mammals. Hum. Gene Ther. 16, 401–407. doi:10.1089/hum.2005.16.401
- Foulongne, V., Sauvage, V., Hebert, C., Dereure, O., Cheval, J., Gouilh, M.A., Pariente, K., Segondy, M., Burguière, A., Manuguerra, J.-C., Caro, V., Eloit, M., 2012. Human Skin Microbiota: High Diversity of DNA Viruses Identified on the Human Skin by High Throughput Sequencing. PLoS One 7, e38499. doi:10.1371/journal.pone.0038499
- Francy, D.S., Stelzer, E.A., Bushon, R.N., Brady, A.M.G., Williston, A.G., Riddell, K.R., Borchardt, M.A., Spencer, S.K., Gellner, T.M., 2012. Comparative effectiveness of membrane bioreactors, conventional secondary treatment, and chlorine and UV disinfection to remove microorganisms from municipal wastewaters. Water Res. 46, 4164–78. doi:10.1016/j.watres.2012.04.044
- Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G., Kelley, J.M., Fritchman, R.D., Weidman, J.F., Small, K. V, Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T.R., Saudek, D.M., Phillips, C.A., Merrick, J.M., Tomb, J.F., Dougherty, B.A., Bott, K.F., Hu, P.C., Lucier, T.S., Peterson, S.N., Smith, H.O., Hutchison, C.A., Venter, J.C., 1995. The minimal gene complement of Mycoplasma genitalium. Science 270, 397–403.
- Fratini, M., Di Bonito, P., La Rosa, G., 2014. Oncogenic Papillomavirus and

Polyomavirus in Water Environments: Is There a Potential for Waterborne Transmission? Food Environ. Virol. 6, 1–12. doi:10.1007/s12560-013-9134-0

- **Freed, E., Martin, M.,** 2013. Fields Virology, Fields Virology. doi:9781451105636
- **Fumian, T.M., da Silva Ribeiro de Andrade, J., Leite, J.P.G., Miagostovich, M.P., Momtaz, Z., Yasmin, R.,** 2016. Norovirus Recombinant Strains Isolated from Gastroenteritis Outbreaks in Southern Brazil, 2004–2011. PLoS One 11, e0145391. doi:10.1371/journal.pone.0145391

G

- **Ganesh, B., Masachessi, G., Mladenova, Z.,** 2014. Animal Picobirnavirus. VirusDisease. doi:10.1007/s13337-014-0207-y
- **Gao, S., Li, D., Zha, E., Zhou, T., Wang, S., Yue, X.,** 2015. Surveillance of hepatitis E virus contamination in shellfish in China. Int. J. Environ. Res. Public Health 12, 2026–36. doi:10.3390/ijerph120202026
- **GenBank database website.** (Accesed 12.05.2017). http://www.ncbi.nlm.nih.gov/genbank
- **Gerba, C.P., Betancourt, W.Q., Kitajima, M.,** 2017. How much reduction of virus is needed for recycled water: A continuous changing need for assessment? Water Res. 108, 25–31. doi:10.1016/j.watres.2016.11.020
- **Gerba, C.P., Goyal, S.M., LaBelle, R.L., Cech, I., Bodgan, G.F.,** 1979. Failure of indicator bacteria to reflect the occurrence of enteroviruses in marine waters. Am. J. Public Health 69, 1116–9.
- **Gibson, K.E., Schwab, K.J., Spencer, S.K., Borchardt, M.A.,** 2012. Measuring and mitigating inhibition during quantitative real time PCR analysis of viral nucleic acid extracts from large-volume environmental water samples. Water Res. 46, 4281–4291. doi:10.1016/j.watres.2012.04.030
- **Ginsberg, H.S.,** 1984. The Adenoviruses. Springer US.
- **Girones, R., Ferrús, M.A., Alonso, J.L., Rodriguez-Manzano, J., Calgua, B., de Abreu Corrêa, A., Hundesa, A., Carratala, A., Bofill-Mas, S.,** 2010. Molecular detection of

pathogens in water – The pros and cons of molecular techniques. Water Res. 44, 4325–4339. doi:10.1016/j.watres.2010.06.030

- Greening, G.E., Hewitt, J., Lewis, G.D., 2002. Evaluation of integrated cell culture-PCR (C-PCR) for virological analysis of environmental samples. J. Appl. Microbiol. 93, 745–750. doi:10.1046/j.1365-2672.2002.01741.x
- Greninger, A.L., Naccache, S.N., Federman, S., Yu, G., Mbala, P., Bres, V., Stryke, D., Bouquet, J., Somasekar, S., Linnen, J.M., Dodd, R., Mulembakani, P., Schneider, B.S., Muyembe-Tamfum, J.-J., Stramer, S.L., Chiu, C.Y., 2015. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. Genome Med. 7, 99. doi:10.1186/s13073-015-0220-9
- Grodzki, M., Schaeffer, J., Piquet, J.-C., Le Saux, J.-C., Cheve, J., Ollivier, J., Le Pendu, J., Le Guyader, F.S., 2014. Bioaccumulation Efficiency, Tissue Distribution, and Environmental Occurrence of Hepatitis E Virus in Bivalve Shellfish from France. Appl. Environ. Microbiol. 80, 4269–4276. doi:10.1128/AEM.00978-14
- Guerrero-Latorre, L., Carratala, A., Rodriguez-Manzano, J., Calgua, B., Hundesa, A., Girones, R., 2011. Occurrence of water-borne enteric viruses in two settlements based in Eastern Chad: analysis of hepatitis E virus, hepatitis A virus and human adenovirus in water sources. J. Water Health 9, 515–24. doi:10.2166/wh.2011.126
- Guimarães, F.R., Ferreira, F.F.M., Vieira, C.B., Fumian, T.M., Shubo, T., Leite, J.P.G., Miagostovich, M.P., 2008. Molecular detection of human astrovirus in an urban sewage treatment plant in Rio de Janeiro, Brazil. Mem. Inst. Oswaldo Cruz 103, 819–823. doi:10.1590/S0074-02762008000800013
- Guiu, R., 2000. Anàlisi de la tendència del consum d ' aigua a Catalunya i marges d ' estalvi 21–33.
- Guix, S., Caballero, S., Villena, C., Bartolomé, R., Latorre, C., Rabella, N., Simó, M., Bosch, A., Pintó, R.M., 2002. Molecular epidemiology of astrovirus infection in Barcelona, Spain. J. Clin. Microbiol. 40, 133–9. doi:10.1128/JCM.40.1.133-139.2002

- Guo, Z., Huang, J., Shi, G., Su, C., Niu, J.J., 2014. A food-borne outbreak of gastroenteritis caused by norovirus GII in a university located in Xiamen City, China. Int. J. Infect. Dis. 28, 101–106. doi:10.1016/j.ijid.2014.06.022
- Gurría, A., 2012. OECD Environmental Outlook to 2050: The Consequences of Inaction - Key Facts and Figures [WWW Document]. Eur. World. doi:ISBN 978-92-64-122161

H

- Häfliger, D., Hübner, P., Lüthy, J., 2000. Outbreak of viral gastroenteritis due to sewage-contaminated drinking water. Int. J. Food Microbiol. 54, 123–126. doi:10.1016/S0168-1605(99)00176-2
- Hall, R.J., Wang, J., Todd, A.K., Bissielo, A.B., Yen, S., Strydom, H., Moore, N.E., Ren, X., Huang, Q.S., Carter, P.E., Peacey, M., 2014. Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. J. Virol. Methods 195, 194–204. doi:10.1016/j.jviromet.2013.08.035
- Halliday, M.L., Kang, L.Y., Zhou, T.K., Hu, M.D., Pan, Q.C., Fu, T.Y., Huang, Y.S., Hu, S.L., 1991. An epidemic of hepatitis A attributable to the ingestion of raw clams in Shanghai, China. J. Infect. Dis. 164, 852–9.
- Hamza, I.A., Jurzik, L., Stang, A., Sure, K., Überla, K., Wilhelm, M., 2009. Detection of human viruses in rivers of a densly-populated area in Germany using a virus adsorption elution method optimized for PCR analyses. Water Res. 43, 2657–2668. doi:10.1016/j.watres.2009.03.020
- Hamza, I.A., Jurzik, L., Uberla, K., Wilhelm, M., 2011. Evaluation of pepper mild mottle virus, human picobirnavirus and Torque teno virus as indicators of fecal contamination in river water. Water Res. 45, 1358–68. doi:10.1016/j.watres.2010.10.021
- Han, T.-H., Kim, C.-H., Park, S.-H., Kim, E.-J., Chung, J.-Y., Hwang, E.-S., 2009. Detection of Human Bocavirus-2 in children with acute Gastroenteritis in South Korea. Arch. Virol. 154, 1923–1927. doi:10.1007/s00705-009-0533-3

- Handelsman, J., 2004. Metagenomics: Application of Genomics to Uncultured Microorganisms. Microbiol. Mol. Biol. Rev. 68, 669–685. doi:10.1128/MMBR.68.4.669-685.2004
- 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. Chem. Biol. 5, R245-9.
- Handley, S.A., Thackray, L.B., Zhao, G., Presti, R., Miller, A.D., Droit, L., Abbink, P., Maxfield, L.F., Kambal, A., Duan, E., Stanley, K., Kramer, J., Macri, S.C., Permar, S.R., Schmitz, J.E., Mansfield, K., Brenchley, J.M., Veazey, R.S., Stappenbeck, T.S., Wang, D., Barouch, D.H., Virgin, H.W., 2012. Pathogenic Simian Immunodeficiency Virus Infection Is Associated with Expansion of the Enteric Virome. Cell 151, 253–266. doi:10.1016/j.cell.2012.09.024
- Hansman, G.S., Oka, T., Li, T.-C., Nishio, O., Noda, M., Takeda, N., 2008a. Detection of human enteric viruses in Japanese clams. J. Food Prot. 71, 1689–95.
- Haramoto, E., Katayama, H., Phanuwan, C., Ohgaki, S., 2008. Quantitative detection of sapoviruses in wastewater and river water in Japan. Lett. Appl. Microbiol. 46, 408–413. doi:10.1111/j.1472-765X.2008.02330.x
- Heegaard, E.D., Brown, K.E., 2002. Human parvovirus B19. Clin. Microbiol. Rev. 15, 485–505. doi:10.1128/CMR.15.3.485-505.2002
- Hellmér, M., Paxéus, N., Magnius, L., Enache, L., Arnholm, B., Johansson, A., Bergström, T., Norder, H., 2014. Detection of pathogenic viruses in sewage provided early warnings of hepatitis A virus and norovirus outbreaks. Appl. Environ. Microbiol. 80, 6771–81. doi:10.1128/AEM.01981-14
- 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, 4523–33.
- Herrero, R., Castellsagué, X., Pawlita, M., Lissowska, J., Kee, F., Balaram, P., Rajkumar, T., Sridhar, H., Rose, B., Pintos, J., Fernández, L., Idris, A., Sánchez, M.J.,

Nieto, A., Talamini, R., Tavani, A., Bosch, F.X., Reidel, U., Snijders, P.J.F., Meijer, C.J.L.M., Viscidi, R., Muñoz, N., Franceschi, S., IARC Multicenter Oral Cancer Study Group, 2003. Human papillomavirus and oral cancer: the International Agency for Research on Cancer multicenter study. J. Natl. Cancer Inst. 95, 1772–83.

- Hjelmsø, M.H., Hellmér, M., Fernandez-Cassi, X., Timoneda, N., Lukjancenko, O., Seidel, M., Elsässer, D., Aarestrup, F.M., Löfström, C., Bofill-Mas, S., Abril, J.F., Girones, R., Schultz, A.C., 2017. Evaluation of Methods for the Concentration and Extraction of Viruses from Sewage in the Context of Metagenomic Sequencing. PLoS One 12, e0170199. doi:10.1371/journal.pone.0170199
- Holm-Hansen, C.C., Midgley, S.E., Fischer, T.K., 2016. Global emergence of enterovirus D68: A systematic review. Lancet Infect. Dis. doi:10.1016/S1473-3099(15)00543-5
- Holtz, L., Wylie, K., Sodergren, E., Jiang, Y., Franz, C., Weinstock, G., Storch, G., Wang, D., 2011. Astrovirus MLB2 Viremia in Febrile Child. Emerg. Infect. Dis. 17, 2050–2. doi:10.3201/eid1711.110496
- Holtz, L.R., Finkbeiner, S.R., Kirkwood, C.D., Wang, D., 2008. Identification of a novel picornavirus related to cosaviruses in a child with acute diarrhea. Virol. J. 5, 159. doi:10.1186/1743-422X-5-159
- Holtz, L.R., Finkbeiner, S.R., Zhao, G., Kirkwood, C.D., Girones, R., Pipas, J.M., Wang, D., 2009a. Klassevirus 1, a previously undescribed member of the family Picornaviridae, is globally widespread. Virol. J. 6, 86. doi:10.1186/1743-422X-6-86
- Hundesa, A., Bofill-Mas, S., Maluquer de Motes, C., Rodriguez-Manzano, J., Bach, A., Casas, M., Girones, R., 2010. Development of a quantitative PCR assay for the quantitation of bovine polyomavirus as a microbial source-tracking tool. J. Virol. Methods 163, 385–389. doi:10.1016/j.jviromet.2009.10.029
- Hundesa, 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. doi:10.1016/j.jviromet.2009.02.006

- 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. Environ. Microbiol. 15, 1428–40. doi:10.1111/j.1462-2920.2012.02836.x
- Hutchison, C.A., Smith, H.O., Pfannkoch, C., Venter, J.C., 2005. Cell-free cloning using 29 DNA polymerase. Proc. Natl. Acad. Sci. 102, 17332–17336. doi:10.1073/pnas.0508809102

I

- Iizuka, S., Oka, T., Tabara, K., Omura, T., Katayama, K., Takeda, N., Noda, M., 2010. Detection of sapoviruses and noroviruses in an outbreak of gastroenteritis linked genetically to shellfish. J. Med. Virol. 82, 1247–1254. doi:10.1002/jmv.21791
- Ijzerman, M.M., Dahling, D.R., Fout, G.S., 1997. A method to remove environmental inhibitors prior to the detection of waterborne enteric viruses by reverse transcription-polymerase chain reaction. J. Virol. Methods 63, 145–53.
- IPCC, 2013. IPCC Fifth Assessment Report (AR5). IPCC.
- Iritani, N., Yamamoto, S.P., Abe, N., Kubo, H., Oka, T., Kaida, A., 2016. Epidemics of GI.2 sapovirus in gastroenteritis outbreaks during 2012–2013 in Osaka City, Japan. J. Med. Virol. 88, 1187–1193. doi:10.1002/jmv.24451
- International Organization for Standardization (2017) ISO 15216-1:2017. Microbiology of the food chain -- Horizontal method for determination of hepatitis A virus and norovirus using real-time RT-PCR -- Part 1: Method for quantification
- International Committee on Taxonomy of Viruses (ICTV). (Accesed 23.03.2017). https://talk.ictvonline.org
- International Human Papillomavirus Reference Center. (Accesed 26.01.2017)

J

- Jacobsen, K.H., Wiersma, S.T., 2010. Hepatitis A virus seroprevalence by age and world region, 1990 and 2005. Vaccine 28, 6653–6657. doi:10.1016/j.vaccine.2010.08.037
- Jiang, S.C., 2006. Human Adenoviruses in Water: Occurrence and Health Implications: A Critical Review †. Environ. Sci. Technol. 40, 7132–7140. doi:10.1021/es060892o
- Jiang, S.C., Chu, W., Hostovesky, T., Johnson, F., 2004. PCR detection of pathogenic viruses in southern California urban rivers. J. Appl. Microbiol. 97, 17–28. doi:10.1111/j.1365-2672.2004.02269.x
- Johne, R., Buck, C.B., Allander, T., Atwood, W.J., Garcea, R.L., Imperiale, M.J., Major, E.O., Ramqvist, T., Norkin, L.C., 2011. Taxonomical developments in the family Polyomaviridae. Arch. Virol. 156, 1627–1634. doi:10.1007/s00705-011-1008-x
- Jones, M.S., Kapoor, A., Lukashov, V. V, Simmonds, P., Hecht, F., Delwart, E., 2005. New DNA viruses identified in patients with acute viral infection syndrome. J. Virol. 79, 8230–6. doi:10.1128/JVI.79.13.8230-8236.2005
- Jones, M.S., Harrach, B., Ganac, R.D., Gozum, M.M.A., Dela Cruz, W.P., Riedel, B., Pan, C., Delwart, E.L., Schnurr, D.P., 2007. New adenovirus species found in a patient presenting with gastroenteritis. J. Virol. 81, 5978–84. doi:10.1128/JVI.02650-06
- Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., Perry, T., Kao, D., Mason, A.L., Madsen, K.L., Wong, G.K.-S., 2016. Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. Front. Microbiol. 7, 459. doi:10.3389/fmicb.2016.00459
- Jurzik, L., Hamza, I.A., Wilhelm, M., 2015. Investigating the Reduction of Human Adenovirus (HAdV) and Human Polyomavirus (HPyV) in a Sewage Treatment Plant with a Polishing Pond as a Tertiary Treatment. Water, Air, Soil Pollut. 226, 284. doi:10.1007/s11270-015-2545-9

K

- Kamar, N., Bendall, R., Legrand-Abravanel, F., Xia, N.-S., Ijaz, S., Izopet, J., Dalton, H.R., 2012. Hepatitis E. Lancet 379, 2477–2488. doi:10.1016/S0140-6736(11)61849-7
- Kapikian, A.Z., Wyatt, R.G., Dolin, R., Thornhill, T.S., Kalica, A.R., Chanock, R.M., 1972. Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis. J. Virol. 10, 1075–81.
- Kaplan, J.E., Goodman, R.A., Schonberger, L.B., Lippy, E.C., Gary, G.W., 1982. Gastroenteritis due to Norwalk virus: an outbreak associated with a municipal water system. J. Infect. Dis. 146, 190–7.
- Kapoor, A., Li, L., Victoria, J., Oderinde, B., Mason, C., Pandey, P., Zaidi, S.Z., Delwart, E., 2009. Multiple novel astrovirus species in human stool. J. Gen. Virol. 90, 2965–2972. doi:10.1099/vir.0.014449-0
- Kapoor, A., Simmonds, P., Slikas, E., Li, L., Bodhidatta, L., Sethabutr, O., Triki, H., Bahri, O., Oderinde, B.S., Baba, M.M., Bukbuk, D.N., Besser, J., Bartkus, J., Delwart, E., 2010. Human bocaviruses are highly diverse, dispersed, recombination prone, and prevalent in enteric infections. J. Infect. Dis. 201, 1633–43. doi:10.1086/652416
- Kapoor, A., Slikas, E., Simmonds, P., Chieochansin, T., Naeem, A., Shaukat, S., Alam, M.M., Sharif, S., Angez, M., Zaidi, S., Delwart, E., 2009. A Newly Identified Bocavirus Species in Human Stool. J. Infect. Dis. 199, 196–200. doi:10.1086/595831
- Kapusinszky, B., Minor, P., Delwart, E., 2012. Nearly constant shedding of diverse enteric viruses by two healthy infants. J. Clin. Microbiol. 50, 3427–34. doi:10.1128/JCM.01589-12
- Karlsson, O.E., Belák, S., Granberg, F., 2013. The effect of preprocessing by sequence-independent, single-primer amplification (SISPA) on metagenomic detection of viruses. Biosecur. Bioterror. 11 Suppl 1, S227-34.

doi:10.1089/bsp.2013.0008

- Karst, S.M., Wobus, C.E., Lay, M., Davidson, J., Virgin, H.W., 2003. STAT1-Dependent Innate Immunity to a Norwalk-Like Virus. Science (80-.). 299, 1575–1578. doi:10.1126/science.1077905
- Katzenelson, E., Fattal, B., Hostovesky, T., 1976. Organic flocculation: an efficient second-step concentration method for the detection of viruses in tap water. Appl. Environ. Microbiol. 32, 638–9.
- Kilianski, A., Haas, J.L., Corriveau, E.J., Liem, A.T., Willis, K.L., Kadavy, D.R., Rosenzweig, C.N., Minot, S.S., 2015. Bacterial and viral identification and differentiation by amplicon sequencing on the MinION nanopore sequencer. Gigascience 4, 12. doi:10.1186/s13742-015-0051-z
- Kim, K., Katayama, H., Kitajima, M., Tohya, Y., Ohgaki, S., 2011. Development of a real-time RT-PCR assay combined with ethidium monoazide treatment for RNA viruses and its application to detect viral RNA after heat exposure. Water Sci. Technol. 63.
- Kirk, M.D., Pires, S.M., Black, R.E., Caipo, M., Crump, J.A., Devleesschauwer, B., Döpfer, D., Fazil, A., Fischer-Walker, C.L., Hald, T., Hall, A.J., Keddy, K.H., Lake, R.J., Lanata, C.F., Torgerson, P.R., Havelaar, A.H., Angulo, F.J., 2015. World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. PLOS Med. 12, e1001921. doi:10.1371/journal.pmed.1001921
- Kirs, M., Smith, D.C., 2007. Multiplex quantitative real-time reverse transcriptase PCR for F+-specific RNA coliphages: a method for use in microbial source tracking. Appl. Environ. Microbiol. 73, 808–14. doi:10.1128/AEM.00399-06
- Kitajima, M., Haramoto, E., Phanuwan, C., Katayama, H., 2011. Prevalence and genetic diversity of Aichi viruses in wastewater and river water in Japan. Appl. Environ. Microbiol. 77, 2184–2187. doi:10.1128/AEM.02328-10
- Kitajima, M., Iker, B.C., Pepper, I.L., Gerba, C.P., 2014. Relative abundance and treatment reduction of viruses during wastewater treatment processes —

Identification of potential viral indicators. Sci. Total Environ. 488, 290–296. doi:10.1016/j.scitotenv.2014.04.087

- **Kobayashi, S., Fujiwara, N., Yasui, Y., Yamashita, T., Hiramatsu, R., Minagawa, H.,** 2012. A foodborne outbreak of sapovirus linked to catered box lunches in Japan. Arch. Virol. 157, 1995–1997. doi:10.1007/s00705-012-1394-8
- **Kokkinos, P., Kozyra, I., Lazic, S., Bouwknegt, M., Rutjes, S., Willems, K., Moloney, R., de Roda Husman, A.M., Kaupke, A., Legaki, E., D'Agostino, M., Cook, N., Rzeżutka, A., Petrovic, T., Vantarakis, A.,** 2012. Harmonised Investigation of the Occurrence of Human Enteric Viruses in the Leafy Green Vegetable Supply Chain in Three European Countries. Food Environ. Virol. 4, 179–191. doi:10.1007/s12560-012-9087-8
- **Koo, E.S., Kim, M.S., Choi, Y.S., Park, K.-S., Jeong, Y.S., Kobayashi, N.,** 2017. Occurrence of novel GII.17 and GII.21 norovirus variants in the coastal environment of South Korea in 2015. PLoS One 12, e0172237. doi:10.1371/journal.pone.0172237
- **Kunin, V., He, S., Warnecke, F., Peterson, S.B., Garcia Martin, H., Haynes, M., Ivanova, N., Blackall, L.L., Breitbart, M., Rohwer, F., McMahon, K.D., Hugenholtz, P.,** 2008. A bacterial metapopulation adapts locally to phage predation despite global dispersal. Genome Res. 18, 293–7. doi:10.1101/gr.6835308
- **Kunze, A., Pei, L., Elsässer, D., Niessner, R., Seidel, M.,** 2015. High performance concentration method for viruses in drinking water. J. Virol. Methods 222, 132–137. doi:10.1016/j.jviromet.2015.06.007
- **Kurtz, J.B., Lee, T.W., Craig, J.W., Reed, S.E.,** 1979. Astrovirus infection in volunteers. J. Med. Virol. 3, 221–30.

L

- **L'Homme, Y., Sansregret, R., Plante-Fortier, É., Lamontagne, A.-M., Ouardani, M., Lacroix, G., Simard, C.,** 2009. Genomic characterization of swine caliciviruses representing a new genus of Caliciviridae. Virus Genes 39, 66–75.

doi:10.1007/s11262-009-0360-3

- La Rosa, G., Fratini, M., Accardi, L., D'Oro, G., Della Libera, S., Muscillo, M., Di Bonito, P., 2013a. Mucosal and Cutaneous Human Papillomaviruses Detected in Raw Sewages. PLoS One 8, e52391. doi:10.1371/journal.pone.0052391
- Lambden, P.R., Caul, E.O., Ashley, C.R., Clarke, I.N., 1993. Sequence and genome organization of a human small round-structured (Norwalk-like) virus. Science 259, 516–9.
- Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25. doi:10.1186/gb-2009-10-3-r25
- Le Cann, P., Ranarijaona, S., Monpoeho, S., Le Guyader, F., Ferré, V., 2004. Quantification of human astroviruses in sewage using real-time RT-PCR. Res. Microbiol. 155, 11–15. doi:10.1016/j.resmic.2003.09.013
- Le Guyader, F.S., Le Saux, J.C., Ambert-Balay, K., Krol, J., Serais, O., Parnaudeau, S., Giraudon, H., Delmas, G., Pommepuy, M., Pothier, P., Atmar, R.L., 2008. Aichi virus, norovirus, astrovirus, enterovirus, and rotavirus involved in clinical cases from a French oyster-related gastroenteritis outbreak. J. Clin. Microbiol. 46, 4011–4017. doi:10.1128/JCM.01044-08
- Lee, N., Chan, M.C.W., Wong, B., Choi, K.W., Sin, W., Lui, G., Chan, P.K.S., Lai, R.W.M., Cockram, C.S., Sung, J.J.Y., Leung, W.K., 2007. Fecal Viral Concentration and Diarrhea in Norovirus Gastroenteritis. Emerg. Infect. Dis. 13, 1399–1401. doi:10.3201/eid1309.061535
- Legrand-Abravanel, F., Mansuy, J.-M., Dubois, M., Kamar, N., Peron, J.-M., Rostaing, L., Izopet, J., 2009. Hepatitis E virus genotype 3 diversity, France. Emerg. Infect. Dis. 15, 110–4. doi:10.3201/eid1501.080296
- Lewis, G.D., Metcalf, T.G., 1988. Polyethylene glycol precipitation for recovery of pathogenic viruses, including hepatitis A virus and human rotavirus, from oyster, water, and sediment samples. Appl. Environ. Microbiol. 54, 1983–8.
- Lewis, H.C., Wichmann, O., Duizer, E., 2010. Transmission routes and risk factors

for autochthonous hepatitis E virus infection in Europe: a systematic review. Epidemiol. Infect. 138, 145. doi:10.1017/S0950268809990847

- Li, L., Kapoor, A., Slikas, B., Bamidele, O.S., Wang, C., Shaukat, S., Masroor, M.A., Wilson, M.L., Ndjango, J.-B.N., Peeters, M., Gross-Camp, N.D., Muller, M.N., Hahn, B.H., Wolfe, N.D., Triki, H., Bartkus, J., Zaidi, S.Z., Delwart, E., 2010. Multiple diverse circoviruses infect farm animals and are commonly found in human and chimpanzee feces. J. Virol. 84, 1674–82. doi:10.1128/JVI.02109-09
- Li, L., Liu, N., Yu, J., Ao, Y., Li, S., Stine, O.C., Duan, Z., 2017. Analysis of Aichi virus and Saffold virus association with pediatric acute gastroenteritis. J. Clin. Virol. 87, 37–42. doi:10.1016/j.jcv.2016.12.003
- Li, T.C., Chijiwa, K., Sera, N., Ishibashi, T., Etoh, Y., Shinohara, Y., Kurata, Y., Ishida, M., Sakamoto, S., Takeda, N., Miyamura, T., 2005. Hepatitis E virus transmission from wild boar meat. Emerg Infect Dis 11, 1958–1960.
- Lim, E.S., Cao, S., Holtz, L.R., Antonio, M., Stine, O.C., Wang, D., 2014. Discovery of rosavirus 2, a novel variant of a rodent-associated picornavirus, in children from The Gambia. Virology 454–455, 25–33. doi:10.1016/j.virol.2014.01.018
- Lodder, W.J., Rutjes, S.A., Takumi, K., de Roda Husman, A.M., 2013. Aichi virus in sewage and surface water, the Netherlands. Emerg. Infect. Dis. 19, 1222–1230. doi:10.3201/eid1908.130312
- Lowther, J.A., Gustar, N.E., Powell, A.L., Hartnell, R.E., Lees, D.N., 2012. Two-Year Systematic Study To Assess Norovirus Contamination in Oysters from Commercial Harvesting Areas in the United Kingdom. Appl. Environ. Microbiol. 78, 5812–5817. doi:10.1128/AEM.01046-12
- Lu, Y., Shen, Y., Warren, W., Walter, R., 2016. Next Generation Sequencing in Aquatic Models, in: Next Generation Sequencing - Advances, Applications and Challenges. InTech. doi:10.5772/61657

M

- Ma, Y., Madupu, R., Karaoz, U., Nossa, C.W., Yang, L., Yooseph, S., Yachimski, P.S., Brodie, E.L., Nelson, K.E., Pei, Z., 2014. Human Papillomavirus Community in Healthy Persons, Defined by Metagenomics Analysis of Human Microbiome Project Shotgun Sequencing Data Sets. J. Virol. 88, 4786–4797. doi:10.1128/JVI.00093-14
- Mäde, D., Trübner, K., Neubert, E., Höhne, M., Johne, R., 2013. Detection and Typing of Norovirus from Frozen Strawberries Involved in a Large-Scale Gastroenteritis Outbreak in Germany. Food Environ. Virol. 5, 162–168. doi:10.1007/s12560-013-9118-0
- Madoni, P., 2011. Protozoa in wastewater treatment processes: A minireview. Ital. J. Zool. 78, 3–11. doi:10.1080/11250000903373797
- Mardis, E.R., 2013. Next-Generation Sequencing Platforms. Annu. Rev. Anal. Chem. 6, 287–303. doi:10.1146/annurev-anchem-062012-092628
- Mardis, E.R., 2008. The impact of next-generation sequencing technology on genetics. Trends Genet. 24, 133–141. doi:10.1016/j.tig.2007.12.007
- Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.-J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X. V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L.I., Jarvie, T.P., Jirage, K.B., Kim, J.-B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M., 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376. doi:10.1038/nature03959
- Marzouk, Y., 1980. Relationship of viruses and indicator bacteria in water and wastewater of Israel. Water Res. 14, 1585–1590. doi:10.1016/0043-

1354(80)90062-7

- Masclaux, F.G., Hotz, P., Friedli, D., Savova-Bianchi, D., Oppliger, A., 2013. High occurrence of hepatitis E virus in samples from wastewater treatment plants in Switzerland and comparison with other enteric viruses. Water Res. 47, 5101–9. doi:10.1016/j.watres.2013.05.050
- Matthijnssens, J., Ciarlet, M., McDonald, S.M., Attoui, H., Bányai, K., Brister, J.R., Buesa, J., Esona, M.D., Estes, M.K., Gentsch, J.R., Iturriza-Gómara, M., Johne, R., Kirkwood, C.D., Martella, V., Mertens, P.P.C., Nakagomi, O., Parreño, V., Rahman, M., Ruggeri, F.M., Saif, L.J., Santos, N., Steyer, A., Taniguchi, K., Patton, J.T., Desselberger, U., Van Ranst, M., 2011. Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). Arch. Virol. 156, 1397–413. doi:10.1007/s00705-011-1006-z
- 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., Moloney, R., D'Agostino, M., de Roda Husman, A.M., von Bonsdorff, C.-H., Rzeżutka, A., Pavlik, I., Petrovic, T., Cook, N., 2013. Tracing enteric viruses in the European berry fruit supply chain. Int. J. Food Microbiol. 167, 177–85. doi:10.1016/j.ijfoodmicro.2013.09.003
- Maunula, L., Miettinen, I.T., von Bonsdorff, C.-H., 2005. Norovirus Outbreaks from Drinking Water. Emerg. Infect. Dis. 11, 1716–1721. doi:10.3201/eid1111.050487
- Mayr, C., Strohe, G., Contzen, M., 2009. Detection of rotavirus in food associated with a gastroenteritis outbreak in a mother and child sanatorium. Int. J. Food Microbiol. 135, 179–82. doi:10.1016/j.ijfoodmicro.2009.08.010
- MetaVir homepage. (Accesed 24.03.2017). http://metavir-meb.univ-bpclermont.fr
- Metzker, M.L., 2010. Sequencing technologies — the next generation. Nat. Rev. Genet. 11, 31–46. doi:10.1038/nrg2626
- Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., Edwards, R., 2008. The metagenomics RAST server – a public resource for the automatic phylogenetic

and functional analysis of metagenomes. BMC Bioinformatics 9, 386. doi:10.1186/1471-2105-9-386

- Meyers, J., Ryndock, E., Conway, M.J., Meyers, C., Robison, R., 2014. Susceptibility of high-risk human papillomavirus type 16 to clinical disinfectants. J. Antimicrob. Chemother. 69, 1546–50. doi:10.1093/jac/dku006
- MG-Rast website. (Accesed 15.03.2017). http://metagenomics.anl.gov/)
- Michen, B., Graule, T., 2010. Isoelectric points of viruses. J. Appl. Microbiol. 109, 388–97. doi:10.1111/j.1365-2672.2010.04663.x
- Monroe, S.S., Jiang, B., Stine, S.E., Koopmans, M., Glass, R.I., 1993. Subgenomic RNA sequence of human astrovirus supports classification of Astroviridae as a new family of RNA viruses. J. Virol. 67, 3611–4.
- Moore, N.E., Wang, J., Hewitt, J., Croucher, D., Williamson, D.A., Paine, S., Yen, S., Greening, G.E., Hall, R.J., 2015. Metagenomic Analysis of Viruses in Feces from Unsolved Outbreaks of Gastroenteritis in Humans. J. Clin. Microbiol. 53, 15–21. doi:10.1128/JCM.02029-14
- Moreas, H., Tsiambas, E., Lazaris, A.C., Nonni, A., Karameris, A., Metaxas, G.E., Armatas, H.E., Patsouris, E., n.d. Impact of HPV detection in colorectal adenocarcinoma: HPV protein and chromogenic in situ hybridization analysis based on tissue microarrays. J. BUON. 19, 91–6.
- Moustafa, A., Xie, C., Kirkness, E., Biggs, W., Wong, E., Turpaz, Y., Bloom, K., Delwart, E., Nelson, K.E., Venter, J.C., Telenti, A., 2017. The blood DNA virome in 8,000 humans. PLOS Pathog. 13, e1006292. doi:10.1371/journal.ppat.1006292
- Müller, L., Rasmussen, L.D., Jensen, T., Schultz, A.C., Kjelsø, C., Barnadas, C., Sigsgaard, K., Larsen, A.R., Jensen, C.W., Jeppesen, S., Uhrbrand, K., Hove, N., Mølbak, K., Ethelberg, S., 2016. Series of Norovirus Outbreaks Caused by Consumption of Green Coral Lettuce, Denmark, April 2016. PLOS Curr. Outbreaks. doi:10.1371/CURRENTS.OUTBREAKS.115761D5D6DE6A8BC7DD4B41F0F5F142
- Muñoz, I., Tomàs, N., Mas, J., García-Reyes, J.F., Molina-Díaz, A., Fernández-Alba, A.R., 2010. Potential chemical and microbiological risks on human health from

urban wastewater reuse in agriculture. Case study of wastewater effluents in Spain. J. Environ. Sci. Heal. Part B 45, 300–309. doi:10.1080/03601231003704648

- Muñoz, N., Bosch, F.X., de Sanjosé, S., Herrero, R., Castellsagué, X., Shah, K. V., Snijders, P.J.F., Meijer, C.J.L.M., International Agency for Research on Cancer Multicenter Cervical Cancer Study Group, 2003. Epidemiologic Classification of Human Papillomavirus Types Associated with Cervical Cancer. N. Engl. J. Med. 348, 518–527. doi:10.1056/NEJMoa021641

N

- Naccache, S.N., Greninger, A.L., Lee, D., Coffey, L.L., Phan, T., Rein-Weston, A., Aronsohn, A., Hackett, J., Delwart, E.L., Chiu, C.Y., Chiu, C.Y., 2013. The perils of pathogen discovery: origin of a novel parvovirus-like hybrid genome traced to nucleic acid extraction spin columns. J. Virol. 87, 11966–77. doi:10.1128/JVI.02323-13
- Nagarajan, N., Pop, M., 2013. Sequence assembly demystified. Nat. Rev. Genet. 14, 157–167. doi:10.1038/nrg3367
- Namiki, T., Hachiya, T., Tanaka, H., Sakakibara, Y., 2012. MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. Nucleic Acids Res. 40, e155. doi:10.1093/nar/gks678
- Nath Srivastava, R., Lund, E., 1980. The stability of bovine parvovirus and its possible use as an indicator for the persistence of enteric viruses. Water Res. 14, 1017–1021. doi:10.1016/0043-1354(80)90146-3
- Nenonen, N.P., Hannoun, C., Horal, P., Hernroth, B., Bergström, T., 2008. Tracing of norovirus outbreak strains in mussels collected near sewage effluents. Appl. Environ. Microbiol. 74, 2544–9. doi:10.1128/AEM.02477-07
- Newell, D.G., Koopmans, M., Verhoef, L., Duizer, E., Aidara-Kane, A., Sprong, H., Opsteegh, M., Langelaar, M., Threfall, J., Scheutz, F., van der Giessen, J., Kruse, H., 2010. Food-borne diseases - the challenges of 20 years ago still persist while new ones continue to emerge. Int. J. Food Microbiol. 139 Suppl, S3-15.

doi:10.1016/j.ijfoodmicro.2010.01.021

- Ng, T.F.F., Marine, R., Wang, C., Simmonds, P., Kapusinszky, 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, 12161–75. doi:10.1128/JVI.00869-12
- Ng, T.F.F., Sachsenröder, J., Reuter, G., Knowles, N.J., Delwart, E., Johne, R., 2015. Rabovirus: a proposed new picornavirus genus that is phylogenetically basal to enteroviruses and sapeloviruses. Arch. Virol. 160, 2569–75. doi:10.1007/s00705-015-2523-y
- Nicolay, N., McDermott, R., Kelly, M., Gorby, M., Prendergast, T., Tuite, G., Coughlan, S., McKeown, P., Sayers, G., 2011. Potential role of asymptomatic kitchen food handlers during a food-borne outbreak of norovirus infection, Dublin, Ireland, March 2009. Euro Surveill. 16.
- Nielsen, A., Bottiger, B., Banner, J., Hoffmann, T., Nielsen, L., 2012. Serious Invasive Saffold Virus Infections in Children, 2009. Emerg. Infect. Dis. 18. doi:10.3201/eid1801.110725
- Nielsen, A.C.Y., Gyhrs, M.L., Nielsen, L.P., Pedersen, C., Böttiger, B., 2013. Gastroenteritis and the novel picornaviruses aichi virus, cosavirus, saffold virus, and salivirus in young children. J. Clin. Virol. 57, 239–242. doi:10.1016/j.jcv.2013.03.015
- Niendorf, S., Jacobsen, S., Faber, M., Eis-Hübinger, A.M., Hofmann, J., Zimmermann, O., Höhne, M., Bock, C.T., 2017. Steep rise in norovirus cases and emergence of a new recombinant strain GII.P16-GII.2, Germany, winter 2016. Euro Surveill. 22. doi:10.2807/1560-7917.ES.2017.22.4.30447
- Nieuwenhuijse, D.F., Koopmans, M.P.G., 2017. Metagenomic Sequencing for Surveillance of Food- and Waterborne Viral Diseases. Front. Microbiol. 8. doi:10.3389/fmicb.2017.00230
- Noble, R.T., Allen, S.M., Blackwood, A.D., Chu, W., Jiang, S.C., Lovelace, G.L., Sobsey, M.D., Stewart, J.R., Wait, D.A., 2003. Use of viral pathogens and indicators

to differentiate between human and non-human fecal contamination in a microbial source tracking comparison study. J. Water Health 1, 195–207.

- Noble, R.T., Fuhrman, J.A., 2001. Enteroviruses detected by reverse transcriptase polymerase chain reaction from the coastal waters of Santa Monica Bay, California: low correlation to bacterial indicator levels. Hydrobiologia 460, 175–184. doi:10.1023/A:1013121416891
- Norja, P., Ubillos, I., Templeton, K., Simmonds, P., 2007. No evidence for an association between infections with WU and KI polyomaviruses and respiratory disease. J. Clin. Virol. 40, 307–311. doi:10.1016/j.jcv.2007.09.008
- Nurk, S., Meleshko, D., Korobeynikov, A., Pevzner, P.A., 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res. gr.213959.116. doi:10.1101/gr.213959.116

O

- O'Ryan, M., 2007. Rotarix™ (RIX4414): an oral human rotavirus vaccine. Expert Rev. Vaccines 6, 11–19. doi:10.1586/14760584.6.1.11
- Oberste, M.S., Gotuzzo, E., Blair, P., Nix, W.A., Ksiazek, T.G., Comer, J.A., Rollin, P., Goldsmith, C.S., Olson, J., Kochel, T.J., 2009. Human febrile illness caused by encephalomyocarditis virus infection, Peru. Emerg. Infect. Dis. 15, 640–6. doi:10.3201/eid1504.081428
- Ogorzaly, L., Cauchie, H.-M., Penny, C., Perrin, A., Gantzer, C., Bertrand, I., 2013. Two-day detection of infectious enteric and non-enteric adenoviruses by improved ICC-qPCR. Appl. Microbiol. Biotechnol. 97, 4159–66. doi:10.1007/s00253-013-4782-4
- Ogorzaly, L., Tissier, A., Bertrand, I., Maul, A., Gantzer, C., 2009. Relationship between F-specific RNA phage genogroups, faecal pollution indicators and human adenoviruses in river water. Water Res. 43, 1257–1264. doi:10.1016/j.watres.2008.12.011
- 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 Environ. Virol. doi:10.1007/s12560-015-9194-4

- Oishi, I., Yamazaki, K., Kimoto, T., Minekawa, Y., Utagawa, E., Yamazaki, S., Inouye, S., Grohmann, G.S., Monroe, S.S., Stine, S.E., Carcamo, C., Ando, T., Glass, R.I., 1994. A Large Outbreak of Acute Gastroenteritis Associated with Astrovirus among Students and Teachers in Osaka, Japan. J. Infect. Dis. 170, 439–443. doi:10.1093/infdis/170.2.439
- Oka, T., Mori, K., Iritani, N., Harada, S., Ueki, Y., Iizuka, S., Mise, K., Murakami, K., Wakita, T., Katayama, K., 2012. Human sapovirus classification based on complete capsid nucleotide sequences. Arch. Virol. 157, 349–352. doi:10.1007/s00705-011-1161-2
- Oller, I., Malato, S., Sánchez-Pérez, J.A., 2011. Combination of Advanced Oxidation Processes and biological treatments for wastewater decontamination—A review. Sci. Total Environ. 409, 4141–4166. doi:10.1016/j.scitotenv.2010.08.061
- Ottoson, J., Hansen, A., Björlenius, B., Norder, H., Stenström, T.A., 2006. Removal of viruses, parasitic protozoa and microbial indicators in conventional and membrane processes in a wastewater pilot plant. Water Res. 40, 1449–57. doi:10.1016/j.watres.2006.01.039
- Oude Munnink, B.B., Jazaeri Farsani, S.M., Deijs, M., Jonkers, J., Verhoeven, J.T.P., Ieven, M., Goossens, H., de Jong, M.D., Berkhout, B., Loens, K., Kellam, P., Bakker, M., Canuti, M., Cotten, M., van der Hoek, L., 2013. Autologous Antibody Capture to Enrich Immunogenic Viruses for Viral Discovery. PLoS One 8, e78454. doi:10.1371/journal.pone.0078454

P

- Painter, J.A., Hoekstra, R.M., Ayers, T., Tauxe, R. V., Braden, C.R., Angulo, F.J., Griffin, P.M., 2013. Attribution of Foodborne Illnesses, Hospitalizations, and Deaths to Food Commodities by using Outbreak Data, United States, 1998–2008.

Emerg. Infect. Dis. 19, 407–415. doi:10.3201/eid1903.111866

- Parashar, U.D., Bresee, J.S., Gentsch, J.R., Glass, R.I., 1998. Rotavirus. Emerg. Infect. Dis. 4, 561–70. doi:10.3201/eid0404.980406
- Parshionikar, S., Laseke, I., Fout, G.S., 2010. Use of Propidium Monoazide in Reverse Transcriptase PCR To Distinguish between Infectious and Noninfectious Enteric Viruses in Water Samples. Appl. Environ. Microbiol. 76, 4318–4326. doi:10.1128/AEM.02800-09
- Parsley, L.C., Consuegra, E.J., Thomas, S.J., Bhavsar, J., Land, A.M., Bhuiyan, N.N., Mazher, M.A., Waters, R.J., Wommack, K.E., Harper, W.F., Liles, M.R., 2010. Census of the viral metagenome within an activated sludge microbial assemblage. Appl. Environ. Microbiol. 76, 2673–7. doi:10.1128/AEM.02520-09
- Patel, M.M., Widdowson, M.-A., Glass, R.I., Akazawa, K., Vinjé, J., Parashar, U.D., 2008. Systematic Literature Review of Role of Noroviruses in Sporadic Gastroenteritis. Emerg. Infect. Dis. 14, 1224–1231. doi:10.3201/eid1408.071114
- Paul, J.H., Jiang, S.C., Rose, J.B., 1991. Concentration of viruses and dissolved DNA from aquatic environments by vortex flow filtration. Appl. Environ. Microbiol. 57, 2197–204.
- Payment, P., Fortin, S., Trudel, M., 1984. Ferric chloride flocculation for nonflocculating beef extract preparations. Appl. Environ. Microbiol. 47, 591–2.
- Pecson, B.M., Martin, L.V., Kohn, T., 2009. Quantitative PCR for determining the infectivity of bacteriophage MS2 upon inactivation by heat, UV-B radiation, and singlet oxygen: advantages and limitations of an enzymatic treatment to reduce false-positive results. Appl. Environ. Microbiol. 75, 5544–54. doi:10.1128/AEM.00425-09
- Pei, L., Rieger, M., Lengger, S., Ott, S., Zawadsky, C., Hartmann, N.M., Selinka, H.-C., Tiehm, A., Niessner, R., Seidel, M., 2012. Combination of Crossflow Ultrafiltration, Monolithic Affinity Filtration, and Quantitative Reverse Transcriptase PCR for Rapid Concentration and Quantification of Model Viruses in Water. Environ. Sci. Technol. 120907094014000. doi:10.1021/es302304t

- Pereira, H.G., 1972. Persistent infection by adenoviruses. J. Clin. Pathol. Suppl. (R. Coll. Pathol). 6, 39–42.
- Pereira, H.G., Fialho, A.M., Flewett, T.H., Teixeira, J.M.S., Andrade, Z.P., 1988. NOVEL VIRUSES IN HUMAN FAECES. Lancet 332, 103–104. doi:10.1016/S0140-6736(88)90032-3
- Peretti, A., FitzGerald, P.C., Bliskovsky, V., Buck, C.B., Pastrana, D. V, 2015. Hamburger polyomaviruses. J. Gen. Virol. 96, 833–9. doi:10.1099/vir.0.000033
- Phan, T.G., Dreno, B., da Costa, A.C., Li, L., Orlandi, P., Deng, X., Kapusinszky, B., Siqueira, J., Knol, A.-C., Halary, F., Dantal, J., Alexander, K.A., Pesavento, P.A., Delwart, E., 2016. A new protoparvovirus in human fecal samples and cutaneous T cell lymphomas (mycosis fungoides). Virology 496, 299–305. doi:10.1016/j.virol.2016.06.013
- Phan, T.G., Vo, N.P., Bonkoungou, I.J.O., Kapoor, A., Barro, N., O'Ryan, M., Kapusinszky, B., Wang, C., Delwart, E., 2012. Acute Diarrhea in West African Children: Diverse Enteric Viruses and a Novel Parvovirus Genus. J. Virol. 86, 11024–11030. doi:10.1128/JVI.01427-12
- Phelps, N.B.D., Mor, S.K., Armien, A.G., Batts, W., Goodwin, A.E., Hopper, L., McCann, R., Ng, T.F.F., Puzach, C., Waltzek, T.B., Delwart, E., Winton, J., Goyal, S.M., 2014. Isolation and molecular characterization of a novel picornavirus from baitfish in the USA. PLoS One 9. doi:10.1371/journal.pone.0087593
- Picanço-Junior, O.M., Oliveira, A.L.T., Freire, L.T.M., Brito, R.B., Villa, L.L., Matos, D., n.d. Association between human Papillomavirus and colorectal adenocarcinoma and its influence on tumor staging and degree of cell differentiation. Arq. Bras. Cir. Dig. 27, 172–6.
- Pickett, B.E., Sadat, E.L., Zhang, Y., Noronha, J.M., Squires, R.B., Hunt, V., Liu, M., Kumar, S., Zaremba, S., Gu, Z., Zhou, L., Larson, C.N., Dietrich, J., Klem, E.B., Scheuermann, R.H., 2012. ViPR: an open bioinformatics database and analysis resource for virology research. Nucleic Acids Res. 40, D593-8. doi:10.1093/nar/gkr859
- Picornavirirus Taxonomy Group. (Accesed 20.02.2017).

http://www.picornaviridae.com/).

- Pina, S., Jofre, J., Emerson, S.U., Purcell, R.H., Girones, R., 1998a. Characterization of a strain of infectious hepatitis E virus isolated from sewage in an area where hepatitis E is not endemic. Appl. Environ. Microbiol. 64, 4485–8.
- Pina, S., Puig, M., Lucena, F., Jofre, J., Girones, R., 1998b. Viral pollution in the environment and in shellfish: human adenovirus detection by PCR as an index of human viruses. Appl. Environ. Microbiol. 64, 3376–82.
- Pintó, R.M., Abad, F.X., Gajardo, R., Bosch, A., 1996. Detection of infectious astroviruses in water. Appl. Environ. Microbiol. 62, 1811–1813.
- Plyusnin, A., Beaty, B.J., Elliott, R.M., Goldbach, R., Kormelink, R., Lundkvist, A., Schmaljohn, C.S., Tesh, R.B., 2011. Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses, vol. 9, in: Edited by King, AMQ, Adams, J. & Lefkowitz, E.. Amsterdam: Elsevier. pp. 735–741.
- Polo, D., Schaeffer, J., Fournet, N., Le Saux, J.-C., Parnaudeau, S., McLeod, C., Le Guyader, F.S., 2016. Digital PCR for Quantifying Norovirus in Oysters Implicated in Outbreaks, France. Emerg. Infect. Dis. 22, 2189–2191. doi:10.3201/eid2212.160841
- Popgeorgiev, N., Temmam, S., Raoult, D., Desnues, C., 2013. Describing the Silent Human Virome with an Emphasis on Giant Viruses. Intervirology 56, 395–412. doi:10.1159/000354561
- Poretsky, R., Rodriguez-R, L.M., Luo, C., Tsementzi, D., Konstantinidis, K.T., 2014. Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. PLoS One 9, e93827. doi:10.1371/journal.pone.0093827
- Prado, T., Silva, D.M., Guilayn, W.C., Rose, T.L., Gaspar, A.M.C., Miagostovich, M.P., 2011. Quantification and molecular characterization of enteric viruses detected in effluents from two hospital wastewater treatment plants. Water Res. 45, 1287–1297. doi:10.1016/j.watres.2010.10.012
- Prez, V.E., Gil, P.I., Temprana, C.F., Cuadrado, P.R., Martínez, L.C., Giordano, M.O.,

Masachessi, G., Isa, M.B., Ré, V.E., Paván, J.V., Nates, S.V., Barril, P.A., 2015. Quantification of human infection risk caused by rotavirus in surface waters from Córdoba, Argentina. Sci. Total Environ. 538, 220–229. doi:10.1016/j.scitotenv.2015.08.041

- Puig, M., Jofre, J., Lucena, F., Allard, A., Wadell, G., Girones, R., 1994. Detection of adenoviruses and enteroviruses in polluted waters by nested PCR amplification. Appl. Environ. Microbiol. 60, 2963–70.
- Pusch, D., Oh, D.-Y., Wolf, S., Dumke, R., Schröter-Bobsin, U., Höhne, M., Röske, I., Schreier, E., 2005. Detection of enteric viruses and bacterial indicators in German environmental waters. Arch. Virol. 150, 929–47. doi:10.1007/s00705-004-0467-8

Q

- Quan, P.L., Wagner, T.A., Briese, T., Torgerson, T.R., Hornig, M., Tashmukhamedova, A., Firth, C., Palacios, G., Baisre-de-Leon, A., Paddock, C.D., Hutchison, S.K., Egholm, M., Zaki, S.R., Goldman, J.E., Ochs, H.D., Lipkin, W.I., 2010. Astrovirus encephalitis in boy with X-linked agammaglobulinemia. Emerg. Infect. Dis. 16, 918–925. doi:10.3201/eid1606.091536
- Quijada, N.M., Fongaro, G., Barardi, C.R.M., Hernández, M., Rodríguez-Lázaro, D., 2016. Propidium Monoazide Integrated with qPCR Enables the Detection and Enumeration of Infectious Enteric RNA and DNA Viruses in Clam and Fermented Sausages. Front. Microbiol. 7, 2008. doi:10.3389/fmicb.2016.02008

R

- Räsänen, S., Lappalainen, S., Kaikkonen, S., Hämäläinen, M., Salminen, M., Vesikari, T., 2010. Mixed viral infections causing acute gastroenteritis in children in a waterborne outbreak. Epidemiol. Infect. 138, 1227–1234. doi:10.1017/S0950268809991671
- RD 140/2003, 2003. Criterios sanitarios de la calidad del agua de consumo

humano.

- RD 1620/2007, 2007. Régimen jurídico de la reutilización de las aguas depuradas
- Reglamento (CE) No 2073/2005, 2005. Criterios microbiológicos aplicables a los productos alimenticios
- Reglamento (UE) No 365/2010, 2010. Modifica modifica el Reglamento (CE) no 2073/2005, relativo a los criterios microbiológicos aplicables a los productos alimenticios, en lo que respecta a las enterobacteriáceas en la leche pasteurizada y otros productos lácteos líquidos pasteurizados y a Listeria monocytogenes en la sal de cocina.
- Reyes, G.R., Kim, J.P., 1991. Sequence-independent, single-primer amplification (SISPA) of complex DNA populations. Mol. Cell. Probes 5, 473–81.
- Reynolds, K.A., Gerba, C.P., Abbaszadegan, M., Pepper, L.L., 2001. ICC/PCR detection of enteroviruses and hepatitis A virus in environmental samples. Can. J. Microbiol. 47, 153–7.
- Riera-Montes, M., Brus Sjölander, K., Allestam, G., Hallin, E., Hedlund, K.-O., Löfdahl, M., 2011. Waterborne norovirus outbreak in a municipal drinking-water supply in Sweden. Epidemiol. Infect. 139, 1928–35. doi:10.1017/S0950268810003146
- Robinson, C.M., Singh, G., Lee, J.Y., Dehghan, S., Rajaiya, J., Liu, E.B., Yousuf, M.A., Betensky, R.A., Jones, M.S., Dyer, D.W., Seto, D., Chodosh, J., 2013. Molecular evolution of human adenoviruses. Sci. Rep. 3, 1812. doi:10.1038/srep01812
- Robinson, C.M., Singh, G., Henquell, C., Walsh, M.P., Peigue-Lafeuille, H., Seto, D., Jones, M.S., Dyer, D.W., Chodosh, J., 2011. Computational analysis and identification of an emergent human adenovirus pathogen implicated in a respiratory fatality. Virology 409, 141–147. doi:10.1016/j.virol.2010.10.020
- Robinson, J.L., Suresh, S., Lee, B.E., 2014. That other EVD: Enterovirus-D68 - what's it all about? Can. J. Infect. Dis. Med. Microbiol. = J. Can. des Mal. Infect. la Microbiol. medicale 25, 296–8.
- Roden, R.B., Lowy, D.R., Schiller, J.T., 1997. Papillomavirus is resistant to

desiccation. J. Infect. Dis. 176, 1076–9.

- Rodríguez-Ezpeleta, N., Hackenberg, M., Aransay, A.M., 2012. Bioinformatics for high throughput sequencing. Springer.
- Rodríguez-Lázaro, D., Cook, N., Ruggeri, F.M., Sellwood, J., Nasser, A., Nascimento, M.S.J., D'Agostino, M., Santos, R., Saiz, J.C., Rzeżutka, A., Bosch, A., Gironés, R., Carducci, A., Muscillo, M., Kovač, K., Diez-Valcarce, M., Vantarakis, A., von Bonsdorff, C.-H., de Roda Husman, A.M., Hernández, M., van der Poel, W.H.M., 2012. Virus hazards from food, water and other contaminated environments. FEMS Microbiol. Rev. 36, 786–814. doi:10.1111/j.1574-6976.2011.00306.x
- Rodriguez-Manzano, J., Alonso, J.L., Ferrús, M.A., Moreno, Y., Amorós, I., Calgua, B., Hundesa, A., Guerrero-Latorre, L., Carratala, A., Rusiñol, M., Girones, R., 2012. Standard and new faecal indicators and pathogens in sewage treatment plants, microbiological parameters for improving the control of reclaimed water. Water Sci. Technol. 66, 2517–23. doi:10.2166/wst.2012.233
- Rosario, K., Nilsson, C., Lim, Y.W., Ruan, Y., Breitbart, M., 2009a. Metagenomic analysis of viruses in reclaimed water. Environ. Microbiol. 11, 2806–2820. doi:10.1111/j.1462-2920.2009.01964.x
- Royall, E., Locker, N., 2016. Translational Control during Calicivirus Infection. Viruses 8, 104. doi:10.3390/v8040104
- Rusiñol, M., Carratalà, A., Hundesa, A., Bach, A., Kern, A., Vantarakis, A., Girones, R., Bofill-Mas, S., 2013. Description of a novel viral tool to identify and quantify ovine faecal pollution in the environment. Sci. Total Environ. 458–460, 355–360. doi:10.1016/j.scitotenv.2013.04.028
- Ruvoën-Clouet, N., Belliot, G., Le Pendu, J., 2013. Noroviruses and histo-blood groups: the impact of common host genetic polymorphisms on virus transmission and evolution. Rev. Med. Virol. 23, 355–366. doi:10.1002/rmv.1757
- Ryndock, E.J., Meyers, C., 2014. A risk for non-sexual transmission of human papillomavirus? Expert Rev. Anti. Infect. Ther. 12, 1165–1170. doi:10.1586/14787210.2014.959497

- Rzeżutka, A., Cook, N., 2004. Survival of human enteric viruses in the environment and food. FEMS Microbiol. Rev. 28, 441–453. doi:10.1016/j.femsre.2004.02.001

S

- Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U. S. A. 74, 5463–7.
- Sano, D., Amarasiri, M., Hata, A., Watanabe, T., Katayama, H., 2016. Risk management of viral infectious diseases in wastewater reclamation and reuse: Review. Environ. Int. 91, 220–229. doi:10.1016/j.envint.2016.03.001
- Sano, D., Perez-Sautu, U., Guix, S., Pinto, R.M., Miura, T., Okabe, S., Bosch, A., 2011. Quantification and Genotyping of Human Sapoviruses in the Llobregat River Catchment, Spain. Appl. Environ. Microbiol. 77, 1111–1114. doi:10.1128/AEM.01721-10
- Santiago-Rodriguez, T.M., Ly, M., Bonilla, N., Pride, D.T., 2015. The human urine virome in association with urinary tract infections. Front. Microbiol. 6, 14. doi:10.3389/fmicb.2015.00014
- Santos, N., Hoshino, Y., 2005. Global distribution of rotavirus serotypes/genotypes and its implication for the development and implementation of an effective rotavirus vaccine. Rev. Med. Virol. 15, 29–56. doi:10.1002/rmv.448
- Sarvikivi, E., Roivainen, M., Maunula, L., Niskanen, T., Korhonen, T., Lappalainen, M., Kuusi, M., 2012. Multiple norovirus outbreaks linked to imported frozen raspberries. Epidemiol. Infect. 140, 260–7. doi:10.1017/S0950268811000379
- Sauvage, V., Laperche, S., Cheval, J., Muth, E., Dubois, M., Boizeau, L., Hébert, C., Lionnet, F., Lefrère, J.-J., Eloit, M., 2016. Viral metagenomics applied to blood donors and recipients at high risk for blood-borne infections. Blood Transfus. 14, 400–7. doi:10.2450/2016.0160-15
- Savichtcheva, O., Okabe, S., 2006. Alternative indicators of fecal pollution: Relations with pathogens and conventional indicators, current methodologies for direct pathogen monitoring and future application perspectives. Water Res. 40,

2463–2476. doi:10.1016/j.watres.2006.04.040

- Sboner, A., Mu, X., Greenbaum, D., Auerbach, R.K., Gerstein, M.B., 2011. The real cost of sequencing: higher than you think! Genome Biol. 12, 125. doi:10.1186/gb-2011-12-8-125
- Scallan, E., Hoekstra, R.M., Angulo, F.J., Tauxe, R. V., Widdowson, M.-A., Roy, S.L., Jones, J.L., Griffin, P.M., 2011. Foodborne Illness Acquired in the United States—Major Pathogens. Emerg. Infect. Dis. 17, 7–15. doi:10.3201/eid1701.P11101
- Scheuer, K.A., Oka, T., Hoet, A.E., Gebreyes, W.A., Molla, B.Z., Saif, L.J., Wang, Q., 2013. Prevalence of Porcine Noroviruses, Molecular Characterization of Emerging Porcine Sapoviruses from Finisher Swine in the United States, and Unified Classification Scheme for Sapoviruses. J. Clin. Microbiol. 51, 2344–2353. doi:10.1128/JCM.00865-13
- Schrader, C., Schielke, A., Ellerbroek, L., Johne, R., 2012. PCR inhibitors - occurrence, properties and removal. J. Appl. Microbiol. 113, 1014–1026. doi:10.1111/j.1365-2672.2012.05384.x
- Shan, T., Lan, D., Li, L., Wang, C., Cui, L., Zhang, W., Hua, X., Zhu, C., Zhao, W., Delwart, E., 2011. Genomic Characterization and High Prevalence of Bocaviruses in Swine. PLoS One 6, e17292. doi:10.1371/journal.pone.0017292
- Shendure, J., Ji, H., 2008. Next-generation DNA sequencing. Nat. Biotechnol. 26, 1135–1145. doi:10.1038/nbt1486
- Shieh, Y.C., Wong, C.I., Krantz, J.A., Hsu, F.C., 2008. Detection of naturally occurring enteroviruses in waters using direct RT-PCR and integrated cell culture-RT-PCR. J. Virol. Methods 149, 184–189. doi:10.1016/j.jviromet.2007.12.013
- Siebrasse, E.A., Reyes, A., Lim, E.S., Zhao, G., Mkakosya, R.S., Manary, M.J., Gordon, J.I., Wang, D., 2012. Identification of MW polyomavirus, a novel polyomavirus in human stool. J. Virol. 86, 10321–6. doi:10.1128/JVI.01210-12
- Sinclair, L., Osman, O.A., Bertilsson, S., Eiler, A., 2015. Microbial community composition and diversity via 16S rRNA gene amplicons: evaluating the illumina platform. PLoS One 10, e0116955. doi:10.1371/journal.pone.0116955

- Smits, S.L., Schapendonk, C.M.E., van Beek, J., Vennema, H., Schürch, A.C., Schipper, D., Bodewes, R., Haagmans, B.L., Osterhaus, A.D.M.E., Koopmans, M.P., 2014. New Viruses in Idiopathic Human Diarrhea Cases, the Netherlands. Emerg. Infect. Dis. 20. doi:10.3201/eid2007.140190
- Smits, S.L., van Leeuwen, M., Schapendonk, C.M.E., Schürch, A.C., Bodewes, R., Haagmans, B.L., Osterhaus, A.D.M.E., 2012. Picobirnaviruses in the Human Respiratory Tract. Emerg. Infect. Dis. 18, 1538–1539. doi:10.3201/eid1809.120507
- Sobsey, M.D., Glass, J.S., 1980. Poliovirus concentration from tap water with electropositive adsorbent filters. Appl. Environ. Microbiol. 40, 201–10.
- Spits, C., Le Caignec, C., De Rycke, M., Van Haute, L., Van Steirteghem, A., Liebaers, I., Sermon, K., 2006. Whole-genome multiple displacement amplification from single cells. Nat. Protoc. 1, 1965–1970. doi:10.1038/nprot.2006.326
- Symonds, E.M., Griffin, D.W., Breitbart, M., 2009a. Eukaryotic viruses in wastewater samples from the United States. Appl. Environ. Microbiol. 75, 1402–9. doi:10.1128/AEM.01899-08

T

- Tan, L. Van, van Doorn, H.R., Nghia, H.D.T., Chau, T.T.H., Tu, L.T.P., de Vries, M., Canuti, M., Deijs, M., Jebbink, M.F., Baker, S., Bryant, J.E., Tham, N.T., BKrong, N.T.T.C., Boni, M.F., Loi, T.Q., Phuong, L.T., Verhoeven, J.T.P., Crusat, M., Jeeninga, R.E., Schultsz, C., Chau, N.V.V., Hien, T.T., van der Hoek, L., Farrar, J., de Jong, M.D., 2013. Identification of a new cyclovirus in cerebrospinal fluid of patients with acute central nervous system infections. MBio 4, e00231-13. doi:10.1128/mBio.00231-13
- Tate, J.E., Burton, A.H., Boschi-Pinto, C., Steele, A.D., Duque, J., Parashar, U.D., WHO-coordinated Global Rotavirus Surveillance Network, 2012. 2008 estimate of worldwide rotavirus-associated mortality in children younger than 5 years before the introduction of universal rotavirus vaccination programmes: a systematic

review and meta-analysis. Lancet. Infect. Dis. 12, 136–41. doi:10.1016/S1473-3099(11)70253-5

- ter Waarbeek, H.L.G., Dukers-Muijrers, N.H.T.M., Vennema, H., Hoebe, C.J.P.A., 2010. Waterborne gastroenteritis outbreak at a scouting camp caused by two norovirus genogroups: GI and GII. J. Clin. Virol. 47, 268–72. doi:10.1016/j.jcv.2009.12.002
- Thorner, A.R., Vogels, R., Kaspers, J., Weverling, G.J., Holterman, L., Lemckert, A.A.C., Dilraj, A., McNally, L.M., Jeena, P.M., Jepsen, S., Abbink, P., Nanda, A., Swanson, P.E., Bates, A.T., O'Brien, K.L., Havenga, M.J.E., Goudsmit, J., Barouch, D.H., 2006. Age dependence of adenovirus-specific neutralizing antibody titers in individuals from sub-Saharan Africa. J. Clin. Microbiol. 44, 3781–3. doi:10.1128/JCM.01249-06
- Thorne, L.G., Goodfellow, I.G., 2014. Norovirus gene expression and replication. J. Gen. Virol. 95, 278–291. doi:10.1099/vir.0.059634-0
- Torner, N., Martinez, A., Broner, S., Moreno, A., Camps, N., Domínguez, A., Working Group for the Study of Acute Viral Gastroenteritis Outbreaks in Catalonia, W.G. for the S. of A.V.G.O. in, 2016. Epidemiology of Acute Gastroenteritis Outbreaks Caused by Human Calicivirus (Norovirus and Sapovirus) in Catalonia: A Two Year Prospective Study, 2010-2011. PLoS One 11, e0152503. doi:10.1371/journal.pone.0152503

U

- United Nations, 2013. World Population Prospects: The 2012 Revision. World Popul. Prospect. 2012 Revis. 1. doi:21 February 2014

V

- Väisänen, E., Kuisma, I., Phan, T.G., Delwart, E., Lappalainen, M., Tarkka, E., Hedman, K., Söderlund-Venermo, M., 2014a. Bufavirus in feces of patients with gastroenteritis, Finland. Emerg. Infect. Dis. 20, 1077–80.

doi:10.3201/eid2006.131674

- Van der Poel, W.H.M., Meng, X.-J., Koenig, M., Harrison, T.J., Oliveira-Filho, E.F., Johne, R., Purdy, M.A., Ulrich, R.G., Simmonds, P., Jameel, S., Okamoto, H., Smith, D.B., Izopet, J., 2016. Proposed reference sequences for hepatitis E virus subtypes. J. Gen. Virol. 97, 537–542. doi:10.1099/jgv.0.000393
- Vaughan, G., Goncalves Rossi, L.M., Forbi, J.C., de Paula, V.S., Purdy, M.A., Xia, G., Khudyakov, Y.E., 2014. Hepatitis A virus: host interactions, molecular epidemiology and evolution. Infect. Genet. Evol. 21, 227–43. doi:10.1016/j.meegid.2013.10.023
- Vázquez-Castellanos, J.F., García-López, R., Pérez-Brocal, V., Pignatelli, M., Moya, A., 2014. Comparison of different assembly and annotation tools on analysis of simulated viral metagenomic communities in the gut. BMC Genomics 15, 37. doi:10.1186/1471-2164-15-37
- Vega, E., Barclay, L., Gregoricus, N., Shirley, S.H., Lee, D., Vinjé, J., 2014. Genotypic and epidemiologic trends of norovirus outbreaks in the United States, 2009 to 2013. J. Clin. Microbiol. 52, 147–55. doi:10.1128/JCM.02680-13
- Vega, E., Barclay, L., Gregoricus, N., Williams, K., Lee, D., Vinjé, J., 2011. Novel surveillance network for norovirus gastroenteritis outbreaks, United States. Emerg. Infect. Dis. 17, 1389–1395. doi:10.3201/eid1708.101837
- Verhoef, L., Hewitt, J., Barclay, L., Ahmed, S.M., Lake, R., Hall, A.J., Lopman, B., Kroneman, A., Vennema, H., Vinjé, J., Koopmans, M., 2015. Norovirus genotype profiles associated with foodborne transmission, 1999-2012. Emerg. Infect. Dis. 21, 592–9. doi:10.3201/eid2104.141073
- Vesikari, T., Matson, D.O., Dennehy, P., Van Damme, P., Santosham, M., Rodriguez, Z., Dallas, M.J., Heyse, J.F., Goveia, M.G., Black, S.B., Shinefield, H.R., Christie, C.D.C., Ylitalo, S., Itzler, R.F., Coia, M.L., Onorato, M.T., Adeyi, B.A., Marshall, G.S., Gothefors, L., Campens, D., Karvonen, A., Watt, J.P., O'Brien, K.L., DiNubile, M.J., Clark, H.F., Boslego, J.W., Offit, P.A., Heaton, P.M., Rotavirus Efficacy and Safety Trial (REST) Study Team, 2006. Safety and Efficacy of a Pentavalent Human–Bovine

(WC3) Reassortant Rotavirus Vaccine. N. Engl. J. Med. 354, 23–33. doi:10.1056/NEJMoa052664

- Victor, R., Kotter, R., O'Brien, G., Mitropoulos, M., Panayi, G., 2008. WHO Guidelines for the Safe Use of Wastewater, Excreta and Greywater, Volumes 1–4. Int. J. Environ. Stud. 65, 157–176. doi:10.1080/00207230701846598
- Vilaginès, P., Sarrette, B., Husson, G., Vilaginès, R., 1993. Glass Wool for Virus Concentration at Ambient Water pH Level. Water Sci. Technol. 27.
- Vipr: The Virus Pathogen Database and Analysis Resource. (Accesed 12.1.2017). www.ViPRbrc.org
- Viral division RefSeq NCBI. (Accesed 15.04.2017). https://www.ncbi.nlm.nih.gov/genome/viruses

W

- Wakuda, M., Pongsuwanna, Y., Taniguchi, K., 2005. Complete nucleotide sequences of two RNA segments of human picobirnavirus. J. Virol. Methods 126, 165–169. doi:10.1016/j.jviromet.2005.02.010
- Wallis, C., Melnick, J.L., 1967. Concentration of enteroviruses on membrane filters. J. Virol. 1, 472–7.
- Walsh, M.P., Seto, J., Liu, E.B., Dehghan, S., Hudson, N.R., Lukashev, A.N., Ivanova, O., Chodosh, J., Dyer, D.W., Jones, M.S., Seto, D., 2011. Computational analysis of two species C human adenoviruses provides evidence of a novel virus. J. Clin. Microbiol. 49, 3482–90. doi:10.1128/JCM.00156-11
- Walter, J.E., Briggs, J., Guerrero, M.L., Matson, D.O., Pickering, L.K., Ruiz-Palacios, G., Berke, T., Mitchell, D.K., 2001. Molecular characterization of a novel recombinant strain of human astrovirus associated with gastroenteritis in children. Arch. Virol. 146, 2357–2367. doi:10.1007/s007050170008
- Wang, D., Coscoy, L., Zylberberg, M., Avila, P.C., Boushey, H. a, Ganem, D., DeRisi, J.L., 2002. Microarray-based detection and genotyping of viral pathogens. Proc. Natl. Acad. Sci. U. S. A. 99, 15687–92. doi:10.1073/pnas.242579699

- Wang, D., Urisman, A., Liu, Y.T., Springer, M., Ksiazek, T.G., Erdman, D.D., Mardis, E.R., Hickenbotham, M., Magrini, V., Eldred, J., Latreille, J.P., Wilson, R.K., Ganem, D., DeRisi, J.L., 2003. Viral discovery and sequence recovery using DNA microarrays. PLoS Biol. 1. doi:10.1371/journal.pbio.0000002
- Wang, Y., Gonzalez, R., Zhou, H., Li, J., Li, Y., Paranhos-Baccalà, G., Vernet, G., Guo, L., Wang, J., 2011. Detection of human bocavirus 3 in China. Eur. J. Clin. Microbiol. Infect. Dis. 30, 799–805. doi:10.1007/s10096-011-1159-4
- Ward, R.L., Bernstein, D.I., Young, E.C., Sherwood, J.R., Knowlton, D.R., Schiff, G.M., 1986. Human rotavirus studies in volunteers: determination of infectious dose and serological response to infection. J. Infect. Dis. 154, 871–80.
- Werness, B.A., Levine, A.J., Howley, P.M., 1990. Association of human papillomavirus types 16 and 18 E6 proteins with p53. Science 248, 76–9.
- Wheaton, TA, Castle, WS, 2001. High application rates of reclaimed water benefit citrus tree growth and fruit production. HORTSCIENCE 36, 1273–1277.
- White, P.A., 2014. Evolution of norovirus. Clin. Microbiol. Infect. 20, 741–745. doi:10.1111/1469-0691.12746
- WHO, 2013a. The world health report 2013: Research for universal health coverage. World Heal. Organ. Press 146. doi:10.1126/scitranslmed.3006971
- WHO, 2013b. Hepatitis A; Factsheet World Health Organization. (Accesed 10.05.2017). http://www.who.int/mediacentre/factsheets/fs328/en/
- WHO, (2015). Hepatitis E [Online]. (Accesed 23.09.2017). Available at: http://www.who.int/mediacentre/factsheets/fs280/en/
- WHO | Food safety, 2016. . WHO.
- Willcocks, M.M., Carter, M.J., 1993. Identification and sequence determination of the capsid protein gene of human astrovirus serotype 1. FEMS Microbiol. Lett. 114, 1–7.
- Wolf, S., Reetz, J., Hoffmann, K., Gründel, A., Schwarz, B.-A., Hänel, I., Otto, P.H., 2012. Discovery and genetic characterization of novel caliciviruses in German and Dutch poultry. Arch. Virol. 157, 1499–1507. doi:10.1007/s00705-012-1326-7

- Wolf, S., Reetz, J., Otto, P., 2011. Genetic characterization of a novel calicivirus from a chicken. Arch. Virol. 156, 1143–1150. doi:10.1007/s00705-011-0964-5
- Wommack, K.E., Sime-Ngando, T., Winget, D.M., Jamindar, S., Helton, R.R., 2010. Filtration-based methods for the collection of viral concentrates from large water samples 110–117.
- Woods, J.W., Calci, K.R., Marchant-Tambone, J.G., Burkhardt, W., 2016. Detection and molecular characterization of norovirus from oysters implicated in outbreaks in the US. Food Microbiol. 59, 76–84. doi:10.1016/j.fm.2016.05.009
- Wylie, K.M., Mihindukulasuriya, K.A., Sodergren, E., Weinstock, G.M., Storch, G.A., 2012. Sequence Analysis of the Human Virome in Febrile and Afebrile Children. PLoS One 7, e27735. doi:10.1371/journal.pone.0027735
- Wylie, K.M., Mihindukulasuriya, K.A., Zhou, Y., Sodergren, E., Storch, G.A., Weinstock, G.M., 2014. Metagenomic analysis of double-stranded DNA viruses in healthy adults. BMC Biol. 12, 71. doi:10.1186/s12915-014-0071-7
- Wylie, T.N., Wylie, K.M., Herter, B.N., Storch, G.A., 2015. Enhanced virome sequencing using targeted sequence capture. Genome Res. 25, 1910–1920. doi:10.1101/gr.191049.115

Y

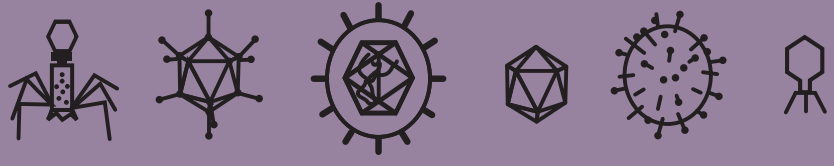
- Yamashita, T., Kobayashi, S., Sakac, K., Nakata, S., Chiba, S., Ishihara, Y., 1991. Isolation of Cytopathic Small Round Viruses with BS-C-l Cells from Patients with Gastroenteritis, The Journal of infectious diseases. doi:10.1093/infdis/164.5.954
- Yang, W., Gu, A.Z., Zeng, S., Li, D., He, M., Shi, H., 2011. Development of a combined immunomagnetic separation and quantitative reverse transcription-PCR assay for sensitive detection of infectious rotavirus in water samples. J. Microbiol. Methods 84, 447–453. doi:10.1016/j.mimet.2011.01.011
- Yazaki, Y., Mizuo, H., Takahashi, M., Nishizawa, T., Sasaki, N., Gotanda, Y., Okamoto, H., 2003. Sporadic acute or fulminant hepatitis E in Hokkaido, Japan, may be food-borne, as suggested by the presence of hepatitis E virus in pig liver as food. J. Gen.

Virol. 84, 2351–2357. doi:10.1099/vir.0.19242-0

Z

- **Zaidi, S.S.Z., Asghar, H., Sharif, S., Alam, M.M.,** 2016. Poliovirus Laboratory Based Surveillance: An Overview, in: Methods in Molecular Biology (Clifton, N.J.). pp. 11–18. doi:10.1007/978-1-4939-3292-4_2
- **Zhang, T., Breitbart, M., Lee, W.H., Run, J.-Q., Wei, C.L., Soh, S.W.L., Hibberd, M.L., Liu, E.T., Rohwer, F., Ruan, Y.,** 2005. RNA Viral Community in Human Feces: Prevalence of Plant Pathogenic Viruses. PLoS Biol. 4, e3. doi:10.1371/journal.pbio.0040003
- **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. doi:10.1016/j.virol.2014.08.025
- **Zhang, Y., Zhu, Z., Yang, W., Ren, J., Tan, X., Wang, Y., Mao, N., Xu, S., Zhu, S., Cui, A., Zhang, Y., Yan, D., Li, Q., Dong, X., Zhang, J., Zhao, Y., Wan, J., Feng, Z., Sun, J., Wang, S., Li, D., Xu, W.,** 2010. An emerging recombinant human enterovirus 71 responsible for the 2008 outbreak of hand foot and mouth disease in Fuyang city of China. Virol. J. 7, 94. doi:10.1186/1743-422X-7-94
- **Zheng, D.-P., Ando, T., Fankhauser, R.L., Beard, R.S., Glass, R.I., Monroe, S.S.,** 2006. Norovirus classification and proposed strain nomenclature. Virology 346, 312–323. doi:10.1016/j.virol.2005.11.015
- **Zhou, N., Lv, D., Wang, S., Lin, X., Bi, Z., Wang, H., Wang, P., Zhang, H., Tao, Z., Hou, P., Song, Y., Xu, A.,** 2016. Continuous detection and genetic diversity of human rotavirus A in sewage in eastern China, 2013–2014. Virol. J. 13, 153. doi:10.1186/s12985-016-0609-0

ANNEXOS



Altres publicacions

A continuació es detallen, per ordre cronològic, altres publicacions no incloses en aquesta tesi però realitzades durant el doctorat.

- 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 Res. Int. 2013, 192089. doi:10.1155/2013/192089

El doctorand va encarregar-se de la revisió bibliogràfica de l'article de revisió i va contribuir en la redacció del manuscrit.

- 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., Bofill-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. doi:10.1016/j.watres.2014.04.013

El doctorand va encarregar-se de l'organització i realització dels mostrejos així com del processament de les mateixes mostres per obtenir els concentrats virals. Posteriorment va realitzar l'anàlisi de les mostres per PCR i qPCR dels NoV GI, NoV GII i HEV.

- Rusiñol, M., Fernandez-Cassi, X., Timoneda, N., Carratalà, A., Abril, J.F., Silvera, C., Figueras, M.J., Gelati, E., Rodó, X., Kay, D., Wyn-Jones, P., Bofill-Mas, S., Girones, R., 2015. Evidence of viral dissemination and seasonality in a Mediterranean river catchment: Implications for water pollution management. J. Environ. Manage. 159, 58–67. doi:10.1016/j.jenvman.2015.05.019

El doctorand va encarregar-se de l'organització i realització dels mostrejos així com del processament de les mateixes mostres per obtenir els concentrats virals. Posteriorment va realitzar l'anàlisi de les mostres per PCR i qPCR dels NoV GI, NoV GII i HEV.

- Sales-Ortells, H., Fernandez-Cassi, X., Timoneda, N., Dürig, W., Girones, R., Medema, G., 2015. Health risks derived from consumption of lettuces irrigated with tertiary effluent containing norovirus. Food Res. Int. 68, 70–77. doi:10.1016/j.foodres.2014.08.018

El doctorand va encarregar-se de la recol·lecció de mostres durant el projecte, la concentració de virus a partir de les mateixes així com de l'anàlisi molecular per determinar norovirus. També va contribuir en la redacció del manuscrit.

- Mayer, R.E., Bofill-Mas, S., Egle, L., Reischer, G.H., Schade, M., Fernandez-Cassi, X., Fuchs, W., Mach, R.L., Lindner, G., Kirschner, A., Gaisbauer, M., Piringer, H., Blaschke, A.P., Girones, R., Zessner, M., Sommer, R., Farnleitner, A.H., 2016. Occurrence of human-associated Bacteroidetes genetic source tracking markers in raw and treated wastewater of municipal and domestic origin and comparison to standard and alternative indicators of faecal pollution. Water Res. 90, 265–276. doi:10.1016/j.watres.2015.12.031

El doctorand va encarregar-se de l'anàlisi molecular per qPCR del virus JCPyV.

- Latif-Eugenín, F., Beaz-Hidalgo, R., Silvera-Simón, C., Fernandez-Cassi, X., Figueras, M.J., 2017. Chlorinated and ultraviolet radiation -treated reclaimed irrigation water is the source of Aeromonas found in vegetables used for human consumption. Environ. Res. 154, 190–195. doi:10.1016/j.envres.2016.12.026

El doctorand va encarregar-se de la recol·lecció, caracterització i enviament de mostres d'aigua regenerada i d'aliments durant el projecte.

Hindawi Publishing Corporation
BioMed Research International
Volume 2013, Article ID 192089, 11 pages
http://dx.doi.org/10.1155/2013/192089

Review Article

Quantification of Human and Animal Viruses to Differentiate the Origin of the Fecal Contamination Present in Environmental Samples

Sílvia Bofill-Mas, Marta Rusiñol, Xavier Fernandez-Cassi, Anna Carratalà, Ayalkibet Hundesa, and Rosina Girones

Laboratory of Viruses Contaminants of Water and Food, Department of Microbiology, Faculty of Biology, Avenida Diagonal 643, 08028 Barcelona, Catalonia, Spain

Correspondence should be addressed to Sílvia Bofill-Mas; sbofill@ub.edu

Received 1 February 2013; Accepted 17 April 2013

Academic Editor: Caroline Rigotto

Many different viruses are excreted by humans and animals and are frequently detected in fecal contaminated waters causing public health concerns. Classical bacterial indicator such as *E. coli* and enterococci could fail to predict the risk for waterborne pathogens such as viruses. Moreover, the presence and levels of bacterial indicators do not always correlate with the presence and concentration of viruses, especially when these indicators are present in low concentrations. Our research group has proposed new viral indicators and methodologies for determining the presence of fecal pollution in environmental samples as well as for tracing the origin of this fecal contamination (microbial source tracking). In this paper, we examine to what extent have these indicators been applied by the scientific community. Recently, quantitative assays for quantification of poultry and ovine viruses have also been described. Overall, quantification by qPCR of human adenoviruses and human polyomavirus JC, porcine adenoviruses, bovine polyomaviruses, chicken/turkey parvoviruses, and ovine polyomaviruses is suggested as a toolbox for the identification of human, porcine, bovine, poultry, and ovine fecal pollution in environmental samples.

1. Fecal Contamination of the Environment

Significant numbers of human microbial pathogens are present in urban sewage and may be considered environmental contaminants. Viruses, along with bacteria and protozoa in the intestine or in urine, are shed and transported through the sewer system. Although most pathogens can be removed by sewage treatment, many are discharged in the effluent and enter receiving waters. Point-source pollution enters the environment at distinct locations, through a direct route of discharge of treated or untreated sewage. Nonpoint sources of contamination are of significant concern with respect to the dissemination of pathogens and their indicators in the water systems. They are generally diffuse and intermittent and may be attributable to the run-off from urban and agricultural areas, leakage from sewers and septic systems, storm water, and sewer overflows [1–3].

Even in highly industrialized countries, viruses that infect humans prevail throughout the environment, causing public health concerns and leading to substantial economic losses. Many orally transmitted viruses produce subclinical infection and symptoms in only a small proportion of the population. However, some viruses may give rise to life-threatening conditions, such as acute hepatitis in adults, as well as severe gastroenteritis in small children and the elderly. The development of disease is related to the infective dose of the viral agent, the age, health, immunological and nutritional status of the infected individual (pregnancy, presence of other infections or diseases), and the availability of health care. Human pathogenic viruses in urban wastewater may potentially include human adenoviruses (HAdVs) and human polyomaviruses (HPyVs), which are detected in all geographical areas and throughout the year, and enteroviruses, noroviruses, rotaviruses, astroviruses,

WATER RESEARCH 59 (2014) 119–129

Available online at www.sciencedirect.com

ScienceDirect

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

Application of human and animal viral microbial source tracking tools in fresh and marine waters from five different geographical areas

Marta Rusiñol[a], Xavier Fernandez-Cassi[a], Ayalkibet Hundesa[a], Carmen Vieira[b], Anita Kern[c], Irene Eriksson[d], Panos Ziros[e], David Kay[f], Marize Miagostovich[b], Marta Vargha[c], Annika Allard[d], Apostolos Vantarakis[e], Peter Wyn-Jones[f], Sílvia Bofill-Mas[a], Rosina Girones[a,*]

[a] Laboratory of Virus Contaminants of Water and Food, Department of Microbiology, University of Barcelona, Barcelona, Spain
[b] Laboratory of Comparative and Environmental Virology, Oswaldo Cruz Institute, Rio de Janeiro, Brazil
[c] Department of Water Microbiology, National Institute for Environmental Health, Budapest, Hungary
[d] Department of Virology, Umeå University, Umeå, Sweden
[e] Environmental Microbiology Unit, Department of Public Health, Medical School, University of Patras, Patras, Greece
[f] Institute of Geography and Earth Sciences (IGES), Aberystwyth University, Durham City, United Kingdom

ARTICLE INFO

Article history:
Received 19 January 2014
Received in revised form
24 February 2014
Accepted 6 April 2014
Available online 18 April 2014

Keywords:
Microbial Source Tracking (MST)
Adenovirus
Polyomavirus
River water
Seawater

ABSTRACT

Integrated river basin management planning to mitigate the impacts of economic, demographic and climate change is an important issue for the future protection of water resources. Identifying sources of microbial contamination via the emerging science of Microbial Source Tracking (MST) plays a key role in risk assessment and the design of remediation strategies. Following an 18-month surveillance program within the EU-FP7-funded VIROCLIME project, specific MST tools were used to assess human markers such as adenoviruses (HAdV) and JC polyomaviruses (JCPyV) and porcine and bovine markers such as porcine adenoviruses (PAdV) and bovine polyomaviruses (BPyV) via quantification with real-time PCR to analyze surface water collected from five sites within different climatic zones: the Negro River (Brazil), Glafkos River (Greece), Tisza River (Hungary), Llobregat River (Spain) and Umeälven River (Sweden). The utility of the viral MST tools and the prevalence and abundance of specific human and animal viruses in the five river catchments and adjacent seawater, which is impacted by riverine contributions from the upstream catchments, were examined. In areas where no sanitation systems have been implemented, sewage can directly enter surface waters, and river water exhibited high viral loads; HAdV and JCPyV could be detected at mean concentrations of 10^5 and 10^4 Genome Copies/Liter (GC/L), respectively. In general, river water samples upstream of urban discharges presented lower human viral loads than downstream sampling sites, and those differences appeared to increase with urban populations but decrease in response to high river flow, as the elevated river water volume dilutes microbial loads. During dry

* Corresponding author. Department of Microbiology, University of Barcelona, Avinguda Diagonal 643, 08028 Barcelona, Spain. Tel.: +34 93 4021483; fax: +34 93 4039047.
E-mail address: rgirones@ub.edu (R. Girones).

http://dx.doi.org/10.1016/j.watres.2014.04.013

Journal of Environmental Management 159 (2015) 58–67

Contents lists available at ScienceDirect

Journal of Environmental Management

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

Research paper

Evidence of viral dissemination and seasonality in a Mediterranean river catchment: Implications for water pollution management

Marta Rusiñol [a], Xavier Fernandez-Cassi [a], Natàlia Timoneda [a, b], Anna Carratalà [a], Josep Francesc Abril [b, c], Carolina Silvera [d], Maria José Figueras [d], Emiliano Gelati [e], Xavier Rodó [e], David Kay [f], Peter Wyn-Jones [f], Sílvia Bofill-Mas [a], Rosina Girones [a, *]

[a] Laboratory of Virus Contaminants of Water and Food, Department of Microbiology, University of Barcelona, Barcelona, Catalonia, Spain
[b] Institute of Biomedicine of the University of Barcelona (IBUB), University of Barcelona, Barcelona, Catalonia, Spain
[c] Computational Genomics Laboratory, Department of Genetics, University of Barcelona, Barcelona, Catalonia, Spain
[d] Microbiology Unit, Faculty of Medicine and Health Sciences, IISPV, University Rovira and Virgili, Reus, Catalonia, Spain
[e] Catalan Institute of Climate Sciences (IC3), Barcelona, Catalonia, Spain
[f] Institute of Geography and Earth Sciences (IGES), Aberystwyth University, Aberystwyth, United Kingdom

ARTICLE INFO

Article history:
Received 25 November 2014
Received in revised form
13 May 2015
Accepted 16 May 2015
Available online 2 June 2015

Keywords:
Human adenovirus
Merkel cell polyomavirus
Norovirus
Hepatitis E virus
River water

ABSTRACT

Conventional wastewater treatment does not completely remove and/or inactive viruses; consequently, viruses excreted by the population can be detected in the environment. This study was undertaken to investigate the distribution and seasonality of human viruses and faecal indicator bacteria (FIB) in a river catchment located in a typical Mediterranean climate region and to discuss future trends in relation to climate change. Sample matrices included river water, untreated and treated wastewater from a wastewater treatment plant within the catchment area, and seawater from potentially impacted bathing water. Five viruses were analysed in the study. Human adenovirus (HAdV) and JC polyomavirus (JCPyV) were analysed as indicators of human faecal contamination of human pathogens; both were reported in urban wastewater (mean values of 10^6 and 10^5 GC/L, respectively), river water (10^3 and 10^2 GC/L) and seawater (10^2 and 10^1 GC/L). Human Merkel Cell polyomavirus (MCPyV), which is associated with Merkel Cell carcinoma, was detected in 75% of the raw wastewater samples (31/37) and quantified by a newly developed quantitative polymerase chain reaction (qPCR) assay with mean concentrations of 10^4 GC/L. This virus is related to skin cancer in susceptible individuals and was found in 29% and 18% of river water and seawater samples, respectively. Seasonality was only observed for norovirus genogroup II (NoV GGII), which was more abundant in cold months with levels up to 10^4 GC/L in river water. Human hepatitis E virus (HEV) was detected in 13.5% of the wastewater samples when analysed by nested PCR (nPCR). Secondary biological treatment (i.e., activated sludge) and tertiary sewage disinfection including chlorination, flocculation and UV radiation removed between 2.22 and 4.52 $\log_{10}$ of the viral concentrations. Climate projections for the Mediterranean climate areas and the selected river catchment estimate general warming and changes in precipitation distribution. Persistent decreases in precipitation during summer can lead to a higher presence of human viruses because river and sea water present the highest viral concentrations during warmer months. In a global context, wastewater management will be the key to preventing environmental dispersion of human faecal pathogens in future climate change scenarios.

1. Introduction

Understanding the environmental fate of pathogens is useful for minimising the risk to humans. Human viruses are excreted at high concentrations in faeces and urine and can be transmitted through improperly treated wastewater. As part of the EU-FP7-funded VIROCLIME project, the present study developed a surveillance program centred on a typically Mediterranean climate region: the Llobregat River basin (Catalonia, northeast of Spain).

Water microbiological quality is traditionally defined and

* Corresponding author. Department of Microbiology, University of Barcelona, Avinguda Diagonal 643, 08028 Barcelona, Spain.
E-mail address: rgirones@ub.edu (R. Girones).

http://dx.doi.org/10.1016/j.jenvman.2015.05.019
0301-4797/

Food Research International 68 (2015) 70–77

Contents lists available at ScienceDirect

Food Research International

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

Health risks derived from consumption of lettuces irrigated with tertiary effluent containing norovirus

Helena Sales-Ortells [a,b,*], Xavier Fernandez-Cassi [c], Natàlia Timoneda [c,d], Wiebke Dürig [a], Rosina Girones [c], Gertjan Medema [a,b]

[a] KWR Watercycle Research Institute, P.O. Box 1072, 3430 BB Nieuwegein, The Netherlands
[b] Section Sanitary Engineering, Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands
[c] Department of Microbiology, Faculty of Biology, University of Barcelona, Av. Diagonal 643, Barcelona 08028, Spain
[d] Computational Genomics Laboratory, Department of Genetics, Institute of Biomedicine, University of Barcelona, Av. Diagonal 643, Barcelona 08028, Spain

ARTICLE INFO

Article history:
Received 22 May 2014
Received in revised form 6 August 2014
Accepted 14 August 2014
Available online 23 August 2014

Keywords:
Norovirus
Water reclamation
Crop irrigation
Health risks
Quantitative Microbial Risk Assessment

ABSTRACT

Wastewater is a valuable resource for water-scarce regions, and is becoming increasingly important due to the rising frequency of droughts as a result of climate change. The health risks derived from ingestion of lettuce that has been irrigated with effluent from a wastewater treatment plant (WWTP) in Catalonia (Spain) were estimated following a quantitative microbial risk assessment (QMRA) approach using site-specific data. Norovirus (NoV) was selected for this analysis, since it is the most common cause of acute gastroenteritis outbreaks in Catalonia. Two scenarios, irrigation with secondary and with tertiary effluent, were analysed. An uncertainty analysis was conducted to determine the impact of possible internalization of NoV into edible parts of the lettuce. The mean disease burden for ingestion of lettuce irrigated with secondary and tertiary effluent was 7.8×10^{-4} Disability Adjusted Life Years (DALYs) per person per year (pppy) and 3.9×10^{-4} DALYs pppy, respectively. A sensitivity analysis revealed that the model parameters with higher influence on the probability of disease are the concentration of NoV in the effluent and the consumption of lettuce. In order to decrease the disease burden to the guidance level of 10^{-6} DALYs pppy, the tertiary treatment should be able to achieve a 4.3 log reduction of the concentration of NoV. If internalization of NoV into lettuces occurs, this would require a reduction of 7.6 log. This is the first time that site specific data and virus internalization in crops are incorporated in a QMRA of irrigation of lettuce and its impact is quantified.

1. Introduction

Wastewater has been widely used in the past for irrigation purposes. It is still in use in developing countries due to water scarcity, the associated nutrient value of these waters for crop growth, and economic limitations. In developed countries, the use of treated wastewater is increasingly seen as a way to deal with water scarcity (exacerbated by climate change), as a more economical alternative to inter-basin transfers, and as an environmentally sustainable practice (Drechsel, Scott, Raschid-Sally, Redwood, & Bahri, 2010).

Uses of reclaimed water include irrigation of landscapes, recreational fields, plants' nurseries, or agricultural lands for food crops, amongst others. In Spain, 362.2 Hm^3 of reclaimed water (42.39 Hm^3 in Catalonia) are used annually, corresponding to 10.6% of the total volume of treated wastewater. 71% of it is used for agricultural irrigation (Iglesias, Ortega, Batanero, & Quintas, 2010).

Although domestic wastewater is treated by secondary or tertiary wastewater treatment, reclaimed water can contain infectious pathogens, posing a risk for public health. Wastewater treatment methodologies are used to reduce concentrations of faecal indicators, e.g. faecal coliforms (FC) or *Escherichia coli* (EC), to below certain standards (BOE, 2007). However, wastewater treatment can be considerably less effective in the elimination of enteric pathogens, such as enteric viruses (EV) and protozoa (Montemayor et al., 2008). Whilst concentrations of FC and EC are usually monitored at the wastewater treatment plants (WWTP), EV, which are relatively resistant to treatment technologies, are not (BOE, 2007), and concentrations of faecal indicators below the standards do not imply absence of EV hazards.

The health risks derived from irrigation of fresh produce with reclaimed water have been previously studied for EV (Hamilton, Stagnitti, Premier, Boland, & Hale, 2006; Petterson, Ashbolt, & Sharma, 2001, 2002; Seidu et al., 2008; Shuval, Lampert, & Fattal, 1997; Stine, Song, Choi, & Gerba, 2005). Few studies focused on the norovirus

* Corresponding author at: KWR Watercycle Research Institute, P.O. Box 1072, 3430 BB Nieuwegein, The Netherlands. Tel.: +31 30 6069 649.
E-mail addresses: hsalesortells@gmail.com (H. Sales-Ortells), xaviako@gmail.com (X. Fernandez-Cassi), Natalia.timoneda@gmail.com (N. Timoneda), wiebkeduerig@gmail.com (W. Dürig), rgirones@ub.edu (R. Girones), gertjan.medema@kwrwater.nl (G. Medema).

http://dx.doi.org/10.1016/j.foodres.2014.08.018
0963-9969/

Water Research 90 (2016) 265–276

Contents lists available at ScienceDirect

Water Research

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

Occurrence of human-associated *Bacteroidetes* genetic source tracking markers in raw and treated wastewater of municipal and domestic origin and comparison to standard and alternative indicators of faecal pollution

R.E. Mayer [a, b, 1], S. Bofill-Mas [i], L. Egle [c, e], G.H. Reischer [a, b, 1], M. Schade [h], X. Fernandez-Cassi [i], W. Fuchs [j], R.L. Mach [a, b, 1], G. Lindner [b, d, 1], A. Kirschner [b, d, 1], M. Gaisbauer [f], H. Piringer [g], A.P. Blaschke [b, e, 1], R. Girones [i], M. Zessner [c, e], R. Sommer [b, d, *, 1], A.H. Farnleitner [a, b, 1]

[a] *Institute of Chemical Engineering, Research Division Biotechnology and Microbiology, Research Group Environmental Microbiology and Molecular Ecology, Vienna University of Technology, Gumpendorfer Straße 1a/166-5-2, A-1060, Vienna, Austria*
[b] *InterUniversity Cooperation Centre for Water and Health, Austria*
[c] *Institute for Water Quality Resources and Waste Management, Vienna University of Technology, Karlsplatz 13/226, 1040, Vienna, Austria*
[d] *Medical University Vienna, Institute for Hygiene and Applied Immunology, Water Hygiene, Kinderspitalgasse 15, A-1090, Vienna, Austria*
[e] *Center of Water Resource Systems, Vienna University of Technology, Karlsplatz 13/222, 1040, Vienna, Austria*
[f] *Schreiber-AWATEC Umwelttechnik GmbH, Bergmillergasse 3/1, 1140, Vienna, Austria*
[g] *VRVis Research Center, Donau-City-Strasse 1, 1220, Vienna, Austria*
[h] *Bavarian Environment Agency, Bürgermeister-Ulrich-Straße 160, 86179, Augsburg, Germany*
[i] *Laboratory of Virus Contaminants of Water and Food, Department of Microbiology, Faculty of Biology, University of Barcelona, Av. Diagonal 643, Barcelona, 08028, Spain*
[j] *Department of Environmental Biotechnology at IFA, University of Natural Resources and Life Sciences, Gregor-Mendel-Straße 33, 1180, Vienna, Austria*

ARTICLE INFO

Article history:
Received 22 July 2015
Received in revised form
13 December 2015
Accepted 17 December 2015
Available online 19 December 2015

Keywords:
HF183 TaqMan
BacHUM
AllBac
Escherichia coli
Clostridium perfringens
Enterococci
JC polyomavirus (JCPyV)
Human adenovirus (HAdV)
Bacteriophage thetaiotaomicron

ABSTRACT

This was a detailed investigation of the seasonal occurrence, dynamics, removal and resistance of human-associated genetic *Bacteroidetes* faecal markers (GeBaM) compared with ISO-based standard faecal indicator bacteria (SFIB), human-specific viral faecal markers and one human-associated *Bacteroidetes* phage in raw and treated wastewater of municipal and domestic origin. Characteristics of the selected activated sludge wastewater treatment plants (WWTPs) from Austria and Germany were studied in detail (WWTPs, n = 13, connected populations from 3 to 49000 individuals), supported by volume-proportional automated 24-h sampling and chemical water quality analysis. GeBaM were consistently detected in high concentrations in raw (median $\log_{10}$ 8.6 marker equivalents (ME) 100 ml^{-1}) and biologically treated wastewater samples (median $\log_{10}$ 6.2–6.5 ME 100 ml^{-1}), irrespective of plant size, type and time of the season (n = 53–65). GeBaM, *Escherichia coli*, and enterococci concentrations revealed the same range of statistical variability for raw (multiplicative standard deviations $s^* = 2.3–3.0$) and treated wastewater ($s^* = 3.7–4.5$), with increased variability after treatment. *Clostridium perfringens* spores revealed the lowest variability for raw wastewater ($s^* = 1.5$). In raw wastewater correlations among microbiological parameters were only detectable between GeBaM, *C. perfringens* and JC polyomaviruses. Statistical associations amongst microbial parameters increased during wastewater treatment. **Two plants with advanced treatment were also investigated, revealing a minimum $\log_{10}$ 5.0 (10th percentile) reduction of GeBaM in the activated sludge membrane bioreactor, but no reduction of the genetic markers during UV irradiation (254 nm). This study highlights the potential of human-associated GeBaM to complement wastewater impact monitoring based on the determination of SFIB. In addition,**

* Corresponding author. Medical University Vienna, Institute for Hygiene and Applied Immunology, Water Hygiene, Kinderspitalgasse 15, A-1090, Vienna, Austria.
E-mail address: regina.sommer@meduniwien.ac.at (R. Sommer).
[1] www.waterandhealth.at.

http://dx.doi.org/10.1016/j.watres.2015.12.031

Environmental Research 154 (2017) 190–195

Contents lists available at ScienceDirect

Environmental Research

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

Chlorinated and ultraviolet radiation -treated reclaimed irrigation water is the source of *Aeromonas* found in vegetables used for human consumption

Fadua Latif-Eugenín[a], Roxana Beaz-Hidalgo[a], Carolina Silvera-Simón[a], Xavi Fernandez-Cassi[b], María J. Figueras[a,*]

[a] Unidad de Microbiología, Facultad de Medicina y Ciencias de la Salud, IISPV, Universidad Rovira i Virgili, Reus, Spain
[b] Departamento de Microbiología, Facultad de Biología, Universidad de Barcelona, Barcelona, Spain

ARTICLE INFO

Keywords:
Reclaimed water
Ready to eat vegetables
Irrigation water
Aeromonas

ABSTRACT

Wastewater is increasingly being recognized as a key water resource, and reclaimed water (or treated wastewater) is used for irrigating vegetables destined for human consumption. The aim of the present study was to determine the diversity and prevalence of *Aeromonas* spp. both in reclaimed water used for irrigation and in the three types of vegetables irrigated with that water. Seven of the 11 (63.6%) samples of reclaimed water and all samples of vegetables were positive for the presence of *Aeromonas*. A total of 216 *Aeromonas* isolates were genotyped and corresponded to 132 different strains that after identification by sequencing the *rpoD* gene belonged to 10 different species. The prevalence of the species varied depending on the type of sample. In the secondary treated reclaimed water *A. caviae* and *A. media* dominated (91.4%) while *A. salmonicida*, *A. media*, *A. allosaccharophila* and *A. popoffii* represented 74.0% of the strains in the irrigation water. In vegetables, *A. caviae* (75.0%) was the most common species, among which a strain isolated from lettuce had the same genotype (ERIC pattern) as a strain recovered from the irrigation water. Furthermore, the same genotype of the species *A. sanarellii* was recovered from parsley and tomatoes demonstrating that irrigation water was the source of contamination and confirming the risk for public health.

1. Introduction

Water is an increasingly scarce resource and as a countermeasure wastewater is being treated to produce reclaimed water that can be reused mainly for agricultural irrigation, including fruits and vegetables that are destined for human consumption (Pianetti et al., 2004; Carvalho et al., 2012; Carey et al., 2016). The safety of reclaimed water and food products (shellfish, lettuces, meat etc.) is evaluated using bacteria indicators of fecal pollution (coliforms, *Escherichia coli*, etc.) during the stipulated controls fixed by legislation to determine their sanitary quality and potential health risk (Figueras and Borrego, 2010; Fernandez-Cassi et al., 2016). However, related illness outbreaks still occur worldwide mainly due to the failure of the fecal indicator organisms to predict the presence of pathogens (Figueras and Borrego, 2010). Among the emerging food and water borne pathogens the bacterial of the genus *Aeromonas* are of special interest because their significance in public health is still not clearly understood (Janda and Abbott, 2010; Figueras and Beaz-Hidalgo, 2015; Teunis and Figueras, 2016).

The genus *Aeromonas* consists of Gram negative, oxidase positive bacilli that are considered autochthonous of aquatic environments and are commonly isolated from clinical and environmental samples (Janda and Abbott, 2010; Beaz-Hidalgo and Figueras, 2013; Figueras and Beaz-Hidalgo, 2014, 2015). Several studies have shown that *Aeromonas* species are foodborne and waterborne pathogens of increasing importance (Abeyta et al., 1986; Granum et al., 1998; Altwegg et al., 1991; Demarta et al., 2000; Figueras and Borrego, 2010; Pablos et al., 2011; Khajanchi et al., 2010; Teunis and Figueras, 2016). *Aeromonas* spp. can be readily isolated from treated sewage, reclaimed water, sea water, fresh water and from drinking water distribution systems, where they appear to survive well, to proliferate at low temperatures and to be linked to pipe biofilms where populations may survive at high chlorine levels (Emekdas et al., 2006; Figueras and Borrego, 2010; Jjemba et al., 2010; Khajanchi et al., 2010; Martone-Rocha et al., 2010; Figueira et al., 2011; Igbinosa and Okoh, 2013; Robertson et al., 2014; Al-Jassim et al., 2015). Several species of *Aeromonas* are recognized to be opportunistic pathogens to humans and can affect both immunocompromised and immunocompetent individuals being the most frequent clinical presentation gastroenteritis followed by wound infections and bacteremia (Janda

* Corresponding author.
E-mail address: mariajose.figueras@urv.cat (M.J. Figueras).

http://dx.doi.org/10.1016/j.envres.2016.12.026
Received 23 October 2016; Received in revised form 18 December 2016; Accepted 26 December 2016
Available online 13 January 2017