UNIVERSITAT DE BARCELONA

Optimización de métodos para el estudio del riesgo asociado a la contaminación fecal del agua y alimentos: cuantificación de patógenos y del riesgo de infecciones virales y análisis metagenómico de virus asociados a hepatitis

Eloy Anibal Gonzales Gustavson

UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA I CIÈNCIES DE L'ALIMENTACIÓ

(Eloy Anibal Gonzales Gustavson)

Optimización de métodos para el estudio del riesgo asociado a la contaminación fecal del agua y alimentos: cuantificación de patógenos y del riesgo de infecciones virales y análisis metagenómico de virus asociados a hepatitis

Eloy Anibal Gonzales Gustavson | 2018

UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA I CIÈNCIES DE L'ALIMENTACIÓ

PROGRAMA DE DOCTORADO DE BIOTECNOLOGÍA

Optimización de métodos para el estudio del riesgo asociado a la contaminación fecal del agua y alimentos: cuantificación de patógenos y del riesgo de infecciones virales y análisis metagenómico de virus asociados a hepatitis

Memoria presentada por Eloy Anibal Gonzales Gustavson para optar el título de doctor por la Universitat de Barcelona

Codirectora	Codirector	Autor
Rosina Girones Llop	Miquel Calvo	Eloy Gonzales Gustavson

Eloy Anibal Gonzales Gustavson | 2018

AGRADECIMIENTOS

En estas pequeñas líneas quiero expresar mi profundo agradecimiento al Programa Nacional de Becas y Crédito Educativo del Gobierno Peruano, por haberme dado la oportunidad de continuar con uno de mis sueños, sin este apoyo nada de esto hubiese sido posible.

También quisiera hacer una mención especial a Rosina Girones, directora de la tesis, por confiar en mí y darme la oportunidad de trabajar bajo su dirección; por su paciencia y por su tiempo. Gracias a ello he aprendido y compartido nuevas experiencias, en diversos lugares y con otras personas.

A mí Codirector Miquel Calvo, por las horas que ha invertido en mí. Desde el primer momento que pude conversar con él, también sentí una gran admiración. Gracias por las conversaciones tan interesantes sobre un tema que, a pesar de no ser mi especialidad, siempre me ha generado gran interés, la estadística.

Y esto me lleva a agradecer también a Armando González y a Teresa López Urbina, mis mentores en mi profesión, la medicina veterinaria, quienes me motivaron e iniciaron en la investigación, la estadística y el trabajo de laboratorio.

A Cecilia, por su amor y apoyo moral, ambos hemos compartido una gran experiencia fenomenal en Barcelona, me has hecho madurar muchísimo como persona y sin lugar a duda esto no habría sido posible sin ti.

A Ayalke, por sus consejos y experiencia, siempre ha estado ahí cuando surgían las dudas acerca del trabajo en el laboratorio, por los cafés de las mañanas y las tantas conversaciones que siempre terminaban en política.

También quiero hacer una mención especial a los profesores Rosa Araujo, Josep Abril y Silvia Bofill, con los que he tenido el placer de trabajar, recibir valiosos consejos y experiencias a la hora de desarrollar los trabajos de investigación que forman parte de esta tesis.

A mis compañeros de laboratorio, Marta, Laura, Xavi, Natalia, Sandra, David y Eva. Gracias por compartir sus experiencias profesionales conmigo, siempre las tendré en cuenta.

A Lily, Yesy y Marcelle, por su amistad incondicional, hemos pasado momentos geniales compartiendo y conociendo esta maravillosa ciudad.

Finalmente, agradecer a Aiora, Pablo, Maryury, al infiltrado Jonas y a William, los chicos del laboratorio 10, con quien compartí cervezas y conversaciones amenas.

TABLA DE CONTENIDO

ABREVIATURAS

ADN	Ácido desoxirribonucleico
ARN	Ácido ribonucleico
BVDV	Virus de la diarrea viral bovina
CG	Copias genómicas
FIB	Indicador fecal bacteriano
HAdV	Adenovirus humanos
HAstV	Astrovirus humanos
HAV	Virus de la hepatitis A
HEV	Virus de la hepatitis E
IFA	Ensayo de inmunofluorescencia
DALY	Años de vida ajustados por la discapacidad
NGS	Secuenciación de nueva generación
NoV	Norovirus
NoV GII	Norovirus genogrupo II
PFU	Unidades formadoras de calvas
Pppy	Por persona por año
qPCR	Reacción en cadena de la polimerasa cuantitativa
QMRA	Análisis cuantitativo del riesgo microbiológico
RoV	Rotavirus
RT	Transcriptasa inversa
SaV	Sapovirus
SMF	Floculación orgánica con leche descremada

TCID50	Dosis infectiva 50
TTV	Torque Teno Virus
TTMDV	Torque Teno Midi Virus
TTMV	Torque Teno Mini Virus
WHO	Organización Mundial de la Salud
WWTP	Planta de tratamiento de agua residual

1 INTRODUCCIÓN:

1.1 El agua como recurso

El agua es esencial para el sostenimiento de la vida en la tierra, por lo tanto, la calidad debe ser controlada para garantizar su uso de forma segura; ya sea para consumo directo como agua de bebida, agricultura, actividades recreacionales o su retorno seguro al medio ambiente después la actividad industrial. Facilitar el acceso a agua segura o su retorno correcto al medio ambiente darán como resultado notorios beneficios para la salud pública (Gorchev and Ozolins, 2011). Una gran diversidad de patógenos está presente en el agua contaminada con materia fecal, lo cual representa un riesgo para la salud de la población. Dentro de estos patógenos podemos incluir virus como los norovirus (NoV), los adenovirus humanos (HAdV) y los rotavirus; bacterias como *Escherichia coli* y *Vibrio cholerae;* protozoos como *Cryptosporidium* spp. y *Giardia* spp. y helmintos (Reynolds et al., 2008).

El uso de agua residual tratada en agricultura es un excelente método de reutilización del agua y sus nutrientes, especialmente en la situación mundial actual de escasez de agua y sobrepoblación. Sin embargo, la presencia de microorganismos patógenos en el agua de riego es un riesgo que puede representar la transmisión de enfermedades. La organización mundial de la salud proporciona directivas para el uso seguro de agua residual en la agricultura (WHO, 2006). Una reducción de 6 a 7 logaritmos en la concentración de microorganismos patógenos en el agua de riego se considera como adecuada. Sin embargo, se recomienda la realización de análisis cuantitativo del riesgo microbiológico (QMRA) basado en escenarios de exposición de consumo de vegetales frescos que permitan alcanzar valores inferiores de 10^{-6} años de vida ajustados por la discapacidad (DALY) por persona por año (pppy).

1.2 Riesgos microbiológicos del consumo de agua con contaminación fecal

Se estima que el consumo de agua de bebida no tratada o inadecuadamente tratada es responsable de 842.000 muertes al año en todo el Mundo (WHO, 2014). Estas muertes están relacionadas a una gran variedad de problemas de salud tanto en humanos como en animales, causando un fuerte impacto en la productividad debido a éstas. Solo en los Estados Unidos, el

consumo de agua contaminada ocasiona pérdidas económicas de aproximadamente 20 mil millones de dólares al año (Amini and Kraatz, 2014).

Tradicionalmente la calidad del agua se ha evaluado mediante el uso de microorganismos indicadores. Los organismos indicadores se utilizan para diversos propósitos: como indicadores de contaminación fecal o para evaluar la eficacia de métodos de tratamiento como la filtración o la desinfección. Los organismos indicadores utilizados más comúnmente son los coliformes termotolerantes, *E. coli* y enterococos intestinales. *E. coli* suele ser el organismo de elección más común, sin embargo, debido a su corta supervivencia en el agua y su sensibilidad a procesos de desinfección, sus valores no se correlacionan con la presencia de otros microrganismos patógenos más resistentes como algunos virus y protozoos de interés (Amini and Kraatz, 2014; Bofill-Mas et al., 2013; Gorchev and Ozolins, 2011).

Otros microrganismos han sido propuestos también como indicadores. Los colifagos son utilizados como indicadores porque comparten propiedades con algunos virus. Sin embargo, tampoco ha sido posible demostrar una directa correlación entre los colifagos y los virus de interés como patógenos. Como consecuencia, se ha recurrido al uso directo de virus que afectan humanos y animales con el propósito de buscar un microorganismo indicador de contaminación fecal (Gorchev and Ozolins, 2011). Se han propuesto a los HAdVs como potenciales indicadores de contaminación humana (Pina et al., 1998), los cuales han sido usados en diferentes estudios como indicadores de contaminación fecal humana así como en estudios de "*microbial source tracking*" (Bofill-Mas et al., 2013, 2011; Rusiñol et al., 2014).

1.2.1 Virus de transmisión fecal-oral

Los virus entéricos son los principales causantes de brotes de enfermedades transmitidas por alimentos en Europa (Sanz and Gawlik, 2014). La lista de virus con transmisión fecal-oral incluye NoV, HAdV, astrovirus (HAstV), rotavirus (RoV), sapovirus (SaV), enterovirus (EV), hepatitis A (HAV) y hepatitis E (HEV). Las características generales de los principales patógenos virales de transmisión hídrica se describen en la tabla 1. Una de las principales vías por las cuales el alimento se contamina con estos virus es a través del agua de riego (Todd and Greig, 2015). Sin embargo, es difícil estimar el impacto de esta vía de transmisión debido:

a. La concentración de virus en los alimentos contaminados es baja y extremadamente variable (Petterson et al., 2015).

b. Los métodos de detección tienen baja sensibilidad (WHO, 2016a).
c. La dosis infectiva para virus es muy baja; siendo en muchos casos muy pocos virus suficientes para infectar un individuo. Por ejemplo: para NoV es aproximadamente 18 partículas víricas (Teunis et al., 2008).

Todos estos factores dificultan la identificación del origen de los brotes. Sin embargo, una alternativa para evaluar el impacto de la transmisión a través de agua y alimentos son los estudios de QMRA (Petterson et al., 2001).

Tabla 1: Principales virus transmitidos a través del agua y analizados en esta tesis.

Clasificación taxonómica	Tamaño del virus y genoma *	Virus	Patologías
Adenoviridae Género *Mastadenovirus*	70 - 100 nm; ADN cadena doble	Adenovirus humano A - G	La mayoría causan enfermedades respiratorias, sin embargo, los serotipos 40 y 41 son los responsables de brotes de gastroenteritis en niños (Wold and Horwitz, 2013)
Reoviridae	76 nm; 11 segmentos de ARN cadena doble	Rotavirus	8 grupos identificados de la A a la H. A nivel global los del grupo A son causantes comunes de vómitos y diarrea severa en recién nacidos y niños menores (Estes and Greenberg, 2013)
Caliciviridae Géneros *Norovirus y Sapovirus*	35 - 40 nm; ARN cadena simple	Norovirus y sapovirus	Causa brotes de gastroenteritis autolimitante y de baja mortalidad; suelen observarse complicaciones en pacientes inmunocomprometidos (Oka et al., 2015)
Astroviridae; Género: *Mamastrovirus*	28 a 34 nm; ARN cadena simple	Astrovirus humanos	Estos virus son de distribución mundial y conocidos por causar gastroenteritis autolimitante (Bosch et al., 2014)
Picornaviridae	60 a 70 nm; ARN de cadena simple	Virus de la hepatitis A	Hepatitis aguda de transmisión fecal-oral (Previsani et al., 2004)
		Enterovirus	Gastroenteritis, encefalitis, meningitis o conjuntivitis
		Aichivirus	Gastroenteritis
Hepeviridae	27 a 34 nm; ARN de cadena simple	HEV	Hepatitis aguda de transmisión fecal-oral con reservorios animales (Ricci et al., 2017)

* Todos presentan geometría icosaédrica y carecen de envoltura lipídica

1.2.2 Bacterias y protozoos analizados en esta tesis

Una gran diversidad de bacterias y protozoos pueden estar presentes en el agua de manera natural, hoy en día se está empezando a considerar al agua como un ambiente microbiano viviente. Los problemas surgen cuando se produce una contaminación, principalmente debida a la introducción de materia fecal en el agua.

E. coli es una bacteria gram negativa, anaerobia facultativa, presentes en el intestino delgado de los organismos de sangre caliente. La mayoría de las cepas de esta bacteria son inofensivas, pero algunos serotipos pueden ser causantes de serios problemas de salud en los humanos y animales; entre ellos están la enteropatogénica, enterotoxigénica, la enteroinvasiva, entre otras (Madigan et al., 2014).

H. pylori es una bacteria gram negativa, microaerófila y tolerante al acido que es aislada comúnmente en el estómago y está asociada a cáncer gástrico. Esta bacteria ha sido reportada en muestras de agua ambiental alrededor del mundo (Eusebi et al., 2014) y se ha demostrado que tiene capacidad de sobrevivir en agua clorada donde la enumeración de coliformes indica que el agua es potable (Santiago et al., 2015)

Giardia lamblia es un protozoo flagelado perteneciente al orden *Diplomonadida*, que parasita el tracto digestivo de humanos y otros mamíferos, produciendo una patología denominada giardiasis. *Criptosporidium parvum* es un protista parasito perteneciente al filo *Apicomplexa*. Tanto giardia como criptosporidium son responsables de brotes de gastroenteritis relacionados con el consumo de agua contaminada (Gascón, 2006). Finalmente, *Acanthamoeba* spp. es un protozoo de vida libre considerado como un patógeno oportunista (Marciano-Cabral and Cabral, 2003). Este protozoo es conocido por ser el responsable de permitir la supervivencia de algunas bacterias patógenas, como *Legionella*, en el agua ambiental.

1.3 Otros virus de interés analizados en esta tesis

1.3.1 Anellovirus

El TTV-1 fue el primer miembro de la familia *Anelloviridae* en ser identificado. Este virus fue descubierto en un paciente con hepatitis al cual no se le encontró ningún otro agente etiológico (Nishizawa et al., 1997). Los virus de esta familia son ADN de cadena simple circular, de forma icosaédrica, sin envoltura y de aproximadamente 30 nm de diámetro. Hasta la fecha, se han

podido identificar tres géneros que infectarían a humanos: *Alphatorquevirus* (conocidos como Torque Teno Virus - TTV), *Betatorquevirus* (conocidos como Torque Teno Mini Virus - TTMV) y *Gammatorquevirus* (conocidos como Torque Teno Midi Virus TTMDV) (Biagini et al., 2012). Sin embargo, el rol de estos virus en el desarrollo de hepatitis permanece desconocido (Hsiao et al., 2016; Okamoto, 2009; Spandole et al., 2015). La prevalencia de este virus esta entre 5% y 90% en muestras de sangre de la población general, dependiendo de la región geográfica (Spandole et al., 2015). Además, la diversidad genética dentro de los anellovirus es una de las más grandes dentro del grupo de virus ADN de cadena simple. Hasta la fecha hay descritas 41 especies que infectan al ser humano y que han sido reconocidas por la ICTV (Biagini et al., 2012). Algunos de estos virus, como el TTV 1, 12, 13, 16, SEN virus D y H, han sido considerados como posibles agentes causales de hepatitis. Adicionalmente, se ha propuesto que infecciones mixtas con estos virus, o la combinación de otros microrganismos, puedan estar asociadas al desarrollo de la enfermedad (Bostan, 2013; Kakkola et al., 2008; Kundu et al., 2013; Mi et al., 2014; Okamoto, 2009).

Desafortunadamente, los anellovirus no pueden ser cultivados *in vitro* debido a que no se cuenta con una línea celular capaz de permitirlo. Sin embargo, estos virus tienen una gran capacidad de replicación *in vivo*. Las infecciones con TTV están caracterizadas por la presencia de una viremia que parece prolongarse durante toda la vida en los humanos, con niveles circulantes en sangre que pueden llegar a 10^6 CG/ml (Okamoto, 2009; Spandole et al., 2015). La replicación del TTV parece darse en el hígado y es excretado en altas concentraciones a través de la bilis y posteriormente en las heces (Ohbayashi et al., 2001). Además, estas partículas virales se hayan frecuentemente en aguas residuales e incluso se han propuesto como indicadores de contaminación viral (Griffin et al., 2008). Sin embargo, otros estudios que sugieren que el tropismo de este virus no se centra únicamente al hígado; su replicación podría producirse en medula ósea, nódulos linfáticos, bazo, páncreas, tiroides, musculo, pulmones, riñón y células sanguíneas mononucleares periféricas (Okamoto, 2009; Spandole et al., 2015). Análisis de metagenómica han demostrado que el TTV es un hallazgo común en diversos tipos de muestras (Delwart, 2007; Rosario et al., 2012). Por tal razón, demostrar que este agente es causa de enfermedad puede llegar a ser un trabajo complicado.

1.3.2 GB virus C

El GB virus C (GBV-C), también conocido como pegivirus o virus de la hepatitis G, es reconocido como un virus que afecta a los humanos y está clasificado dentro de la familia *Flaviviridae*;

relacionándolo estructural y epidemiológicamente al virus de la hepatitis C (Chivero and Stapleton, 2015). La mayoría de las infecciones ocasionadas por el GBV-C son aparentemente asintomáticas, transitorias y autolimitantes, con elevaciones ligeras de los niveles de ALT que pueden pasar desapercibidas. Estas infecciones son difíciles de identificar y mucho más difíciles de valorar. Se desconoce cuál es el rol de este virus en el desarrollo de la hepatitis o en cualquier otra enfermedad (Leary and Mushahwar, 2004). Además, es un hallazgo muy común en estudios de metagenómica (Delwart, 2007), esto sugiere que posiblemente no desempeñen ningún papel en el desarrollo de alguna enfermedad, incluyendo la hepatitis (Chivero and Stapleton, 2015).

1.3.3 Virus de la diarrea viral bovina

Este virus, comúnmente conocido por sus siglas en ingles BVDV “Bovine viral diarrea virus” es un virus ARN de cadena simple, envuelto por una membrana lipídica y pertenece a la familia *Flavivivridae*. Este virus es causante de importantes pérdidas económicas a nivel mundial debido a su alta morbilidad y mortalidad en el ganado vacuno (MacLachlan and Dubovi, 2011). Es un virus cultivable y por ese motivo se ha utilizado en esta tesis como representante de los virus con envoltura lipídica para evaluar su detección en el agua.

1.3.4 Bacteriófago MS2

El bacteriofago MS2 es un virus ARN de cadena simple, de simetría icosaédrica y sin envoltura que pertenece a la familia *Leviviridae*. El MS2 infecta a *E. coli* a través del pili sexual por lo que se encuentra entre los denominados colifagos F-RNA. Además, es utilizado comúnmente como control de proceso en el análisis de agua y alimentos (van Duin and Olsthoorn, 2012)

1.4 Cuantificación de patógenos virales en agua

Uno de los principales problemas para determinar la calidad microbiológica del agua es la escasa y variable eficiencia de los métodos de concentración a la hora de identificar y cuantificar los microorganismos. Como es sabido, la evaluación directa de estos en el agua es difícil debido a que se encuentran en muy bajas concentraciones que fluctúan en el espacio (no son homogéneas) y en el tiempo. Cuando se utilizan volúmenes los suficientemente grandes como para que la muestra sea representativa, la mayoría de técnicas con las que se cuenta hoy en día tienen como inconveniente de haber sido diseñadas para microorganismos específicos, y finalmente estos métodos son altamente variables o no se tiene conocimiento de su repetitividad (Amini and Kraatz, 2014).

1.4.1 Técnicas de concentración de microorganismos en el agua

Existen múltiples métodos de concentración de los microorganismos presentes en el agua; su elección depende de múltiples características como el tipo de agua, el volumen y el tipo de microrganismos que se quiere detectar. Tradicionalmente, el método de filtración con membrana es el más usado para detectar bacterias en diferentes tipos de muestras de agua porque es una método rápido y de bajo coste (Amini and Kraatz, 2014). Sin embargo, no existe un consenso claro acerca de que método de concentración utilizar para detectar algunos virus fecales.

Muchos métodos de concentración de virus funcionan a través de procesos de adsorción y elución que son dependientes de las cargas eléctricas originadas por el punto isoeléctrico de las estructuras virales externas (Michen and Graule, 2010). Los virus, al tener una carga negativa neta, pueden adsorberse a membranas, filtros o diferentes matrices. Estos métodos requieren muestras relativamente limpias debido a que es muy fácil que el sistema se pueda obstruir. Además, los porcentajes de recuperación suelen ser muy variables (Albinana-Gimenez et al., 2009; Lambertini et al., 2008; Wyn-Jones et al., 2011). Otros grupos de métodos se basan en procesos de ultrafiltración y para pequeños volúmenes ultracentrifugación o la combinación de ultrafiltración y centrifugación (Brinkman et al., 2013; Prata et al., 2012).

1.4.1.1 Floculación orgánica con leche descremada

Se han ensayado diversos métodos de floculación orgánica; en esta tesis describimos el método que nos ha permitido obtener los mejores resultados la floculación orgánica con leche descremada (SMF de sus siglas en inglés *Skimmed milk Flocculation*). La SMF es un método muy sencillo y de bajo coste que sirve para concentrar virus en todo tipo de muestras de agua. Este método ha sido utilizado para detectar virus en muestras ambientales tales como: agua de rio (Calgua et al., 2013a), agua de mar (Calgua et al., 2008), aguas subterráneas (Bofill-Mas et al., 2011) e incluso en muestras de agua residuales con altos niveles de turbidez (Calgua et al., 2013c). Sin embargo, su eficiencia no ha sido completamente caracterizada.

1.4.2 Técnicas de cuantificación de virus

1.4.2.1 Métodos moleculares

Con respecto a los métodos de detección y cuantificación de patógenos e indicadores, los ensayos moleculares, especialmente los métodos cuantitativos como la reacción en cadena de la polimerasa cuantitativa (qPCR) o los que llevan un paso previo de transcriptasa reversa (qRTPCR),

son los más utilizados para determinar la presencia y cantidad de virus ambientales de origen humano en muestras de agua. Esto es debido a que estos métodos son más sensibles, específicos, versátiles, precisos y fiables al detectar los ácidos nucleicos en las muestras ambientales de agua (Amini and Kraatz, 2014). Otra ventaja de estos métodos es que son capaces de detectar virus que no crecen en líneas celulares. Además, los ácidos nucleicos virales libres: a) se degradan rápidamente (en cuestión de minutos) en el agua residual, por lo que solo se detectarían viriones intactos y b) no se concentran a la misma eficiencia que los virus intactos. Estos factores hacen de los métodos moleculares la opción de elección (Gerba et al., 2017).

1.4.2.2 Cuantificación de partículas víricas infectivas

Otros métodos son los ensayos de infectividad, como las titulaciones virales a través de: inmunofluorescencia (IFA de sus siglas en inglés *immunofluorescence assays*), unidades formadoras de calvas (PFU del inglés *plaque forming units*) o dosis infectiva 50 (TCID50 del inglés *tissue culture infective dose*); son consideradas técnicas de detección tradicionales para cuantificar partículas infectivas. Los inconvenientes de la utilización de estos métodos es que son muy costosos, requieren mucho tiempo e instalaciones complejas para poder desarrollarlos y además es necesario personal altamente cualificado. Así mismo, se reconocen algunos factores que pueden influenciar la habilidad de detectar virus en líneas celulares como el tipo de línea, el números de pases, el tiempo de exposición o la agregación de las partículas víricas, entre otros (Gerba et al., 2017). Sin embargo, la principal ventaja es que estos métodos, en la mayoría de los casos, permiten una determinación cuantitativa de los microorganismos viables (Rames et al., 2016).

1.5 Análisis cuantitativo del riesgo microbiológico

El análisis de riesgo es una herramienta científica que se utiliza para caracterizar la seguridad microbiológica del agua y resulta necesaria para desarrollar estrategias exitosas que permitan gestionar el riesgo de los patógenos en la salud humana. Frecuentemente en los modelos de análisis de riesgo se utiliza la distribución de probabilidad de la variable evaluada. La ventaja de este resultado es que el riesgo es representado a través de una función de distribución de probabilidad en vez de un valor puntual. De esta manera se tienen en consideración todos los posibles escenarios en función de la información de las distribuciones introducidas al modelo. El objetivo del análisis de riesgo es calcular el impacto combinado de la variabilidad y la

incertidumbre en los parámetros del modelo con el fin de determinar una distribución de la incertidumbre que incluya todos los posibles valores del modelo (Vose, 2008).

El QMRA es un método que permite estimar el riesgo potencial de los microrganismos (WHO, 2016a). Las directrices de la Organización mundial de la salud relacionadas con el agua recomiendan el uso de esta herramienta matemática para describir el riesgo y evaluar potenciales estrategias preventivas para su reutilización (Gorchev and Ozolins, 2011).

El enfoque del QMRA combina el conocimiento científico sobre la presencia y naturaleza de los patógenos, su potencial transporte y destino final en el ciclo del agua, las rutas de exposición a los humanos y los efectos sobre la salud que pueden resultar de su exposición. Todo este conocimiento es combinado dentro de una evaluación que permite un manejo objetivo del riesgo de las enfermedades trasmitidas por el agua (WHO, 2016a). Esta es la forma más efectiva para asegurar el uso correcto de agua regenerada en la agricultura, la cual nos permitirá determinar el riesgo de su utilización y qué medidas tomar para disminuir su impacto (WHO, 2006).

El QMRA está compuesto de cuatro pasos: formulación del problema, evaluación de la exposición, evaluación de los efectos en la salud y caracterización del riesgo, los cuales son descritos comúnmente en un diagrama de flujo al que se conoce como modelo conceptual (Vose, 2008).

1.5.1 Formulación del problema

La formulación del problema es una planificación sistemática que identifica el propósito del QMRA. En este paso se define el contexto general del análisis de riesgo: patógenos usados como referencia, vías de exposición, eventos que incrementan el riesgo y el indicador de salud de interés. Normalmente este paso requiere un equipo interdisciplinario que defina el alcance del análisis en función de los siguientes puntos:

1.5.1.1 Identificación del peligro

Debido a que no es posible identificar y cuantificar todos los patógenos en el agua, es necesario elegir patógenos de referencia que sirvan como indicador de los patógenos de interés. Los patógenos de referencia son seleccionados en función de las condiciones locales, incluyendo las vías de exposición, las características de las fuentes de agua, la incidencia y la severidad de las enfermedades transmitidas por el agua.

1.5.1.2 Identificación de las vías de exposición

Se debe definir la vía por la cual los patógenos llegan a la población objetivo, para los propósitos de esta tesis es el agua.

1.5.1.3 Indicador de salud

Se debe establecer el indicador de salud de interés. Dependiendo de los propósitos del ensayo, el indicador de salud puede incluir: infección, enfermedad, secuelas o más comúnmente una medida que combina todos los anteriores que se conoce como los DALYs (WHO, 2016a), cuya definición se indica más adelante.

1.5.2 Evaluación de la exposición

El objetivo de la evaluación de la exposición es estimar la magnitud y la frecuencia de exposición a través de la identificación de las vías de exposición y la duración de los eventos de peligro definidos previamente en la formulación del problema.

La evaluación de la exposición involucra los siguientes pasos:

a. Definir las vías de exposición que fueron previamente formuladas en detalle, incluyendo los puntos de cuantificación de la fuente de patógenos, reducciones (o recontaminaciones) debido a procesos naturales o realizados por el hombre y los mecanismos de exposición, todos estos conforman los escenarios de exposición. Esta aproximación facilita la armonización de la interpretación de los datos y métodos estadísticos que conlleva la reutilización del agua.
b. Cuantificación de cada componente de la vía de exposición usando la mejor evidencia científica disponible y el entendimiento de la variabilidad e incertidumbre esperada que se asocia con cada variable modelada.
c. La caracterización de la exposición que expresa cuantitativamente la magnitud y frecuencia de la exposición para un rango de escenarios previamente definidos en el QMRA.

Como resultado final de la evaluación de la exposición obtendremos la “unidad de dosis” a la cual los individuos estarán expuestos. Esta se suele determinar en función de la concentración de patógenos en el medio de exposición y la cantidad de material ingerido o inhalado por evento (WHO, 2016a).

1.5.3 Evaluación de los efectos en la salud

En este paso se relacionan los datos de impacto en la salud para identificar los peligros de la población en estudio. Dependiendo del indicador de salud requerido para el análisis (identificado durante la formulación del problema), se necesitan considerar los siguientes componentes:

a. Dosis-Respuesta: la aplicación del modelo de dosis-respuesta es la conexión entre la exposición de patógenos y el indicador de salud (sea infección o enfermedad). El modelo debe ser seleccionado en función de la información presente en la literatura para cada estudio en particular. Como resultado de la aplicación del modelo seleccionado, se obtiene la probabilidad de infección diaria en función de la unidad de dosis previamente descrita en la evaluación de la exposición.
b. Probabilidad de enfermedad: no todos los individuos infectados desarrollan signos o síntomas de enfermedad. Cuando se usan modelos de dosis-respuesta que están en función de la infección, es necesario también estimar la probabilidad de enfermedad condicionada a que el individuo esté infectado.
c. Carga de enfermedad: los DALYs son una medida recomendada en las directrices de la WHO para estimar la salud poblacional. Para las enfermedades transmitidas a través del agua, éste incorpora el impacto total de todos los anteriores indicadores mencionados en la población expuesta. La ventaja de utilizar los DALYs es que permite considerar el impacto de la enfermedad en función de la calidad y la cantidad de vida perdida en la población (WHO, 2016a).

1.5.4 Caracterización del riesgo

En este paso, la información proveniente de la evaluación de la exposición y la evaluación de los efectos en la salud se combinan para expresar una estimación cuantitativa del riesgo.

La estimación cuantitativa del riesgo puede ser evaluada de forma determinista o con una perspectiva probabilista. El primero utiliza valores medios o estimaciones puntuales y el método probabilístico incluye distribuciones de probabilidad para describir la variabilidad e incertidumbre que existe al introducir información en el modelo.

Como se mencionó anteriormente en la formulación del problema, existen varios indicadores de salud que permiten caracterizar el riesgo como la probabilidad de infección, enfermedad o los DALYs. Estos pueden ser definidos en diferentes escalas, incluyendo exposiciones diarias y anuales.

La estimación anual del riesgo es utilizada más comúnmente debido a dos ventajas: a) una alta probabilidad de infección podría ser tolerable cuando la exposición es poco frecuente, b) la estimación anual permite observar la variabilidad del riesgo diario lo cual permite en ciertas ocasiones tolerar riesgos más altos (WHO, 2016a).

1.5.4.1 Estimación anual del riesgo

La estimación anual del riesgo, calculada en función de la exposición a múltiples eventos, se calcula en función de la probabilidad de infección (o enfermedad) obtenida para un evento individual (o diario). La fórmula tradicionalmente utilizada es:

$$P_{\inf anual} = 1 - (1 - P_{\inf diaria})^N$$

Donde $P_{\inf anual}$ es la probabilidad de infección (o enfermedad) anual sobre un N número de exposiciones que se producen en el año, usualmente 365, $P_{\inf diaria}$ es la probabilidad de infección diaria (Haas et al., 1999).

Sin embargo, la formula anterior no tiene en consideración la variabilidad en la probabilidad de infección diaria, ya que durante un período de tiempo definido este no es constante. Por lo tanto, una forma más realista para evaluar esta probabilidad es:

$$P_{inf\ anual} = 1 - \prod_{1}^{N}\left(1 - Random(P_{\inf diaria})\right)$$

Donde $Random(P_{\inf diaria})$ es una extracción aleatoria de acuerdo a la distribución de $P_{\inf diaria}$, que se repite N veces y conforma la probabilidad de infección anual de acuerdo a la expresión anterior (Karavarsamis and Hamilton, 2010).

1.5.4.2 Estimación de los DALYs:

Los años de vida ajustados por la discapacidad, de sus siglas en inglés *Disability Adjusted life Years,* es una medida de la carga de enfermedad que agrega el impacto de todos sus efectos. Los DALYs representan un valor que resume la salud de la población e incorpora la severidad y la duración debido a una enfermedad. Los DALYs han sido incorporados como un indicador dentro de las directrices de la WHO con el fin de proveer un peso relativo a las enfermedades en función de su severidad (Havelaar and Melse, 2003).

Un DALY representa la pérdida de un año de vida saludable. Los DALYs se calculan sumando los años de vida perdidos debido a una muerte prematura más los años perdidos debido a la discapacidad que ocasiona la enfermedad. Se representa con la siguiente formula:

$$DALY = YLL + YLD$$

Donde YLL son los años de vida perdidos (del inglés Years of life lost) y YLD son los años de vida con discapacidad por la enfermedad (del inglés Years living with a disability).

Los años de vida perdidos se pueden calcular con la siguiente ecuación:

$$YLL = Nd \times L$$

Donde Nd es el número de muertes y L es los años promedios perdidos debido a la enfermedad.

Los años de vida con discapacidad por la enfermedad se pueden expresar con el siguiente producto:

$$YLD = Dt \times Dw$$

Donde Dt es la duración de la enfermedad y Dw es un indicador del peso de la discapacidad.

Una manera más simplificada para estimar los DALYs causados por un patógeno también puede ser calculada de la siguiente forma:

$$DALY = P_{ill\ annual} \times DBPC \times F_s$$

Donde $P_{ill\ annual}$ es la probabilidad anual de enfermedad, $DBPC$ corresponde a la carga de enfermedad por caso que causa cada enfermedad y F_s es la proporción de la población susceptible a la enfermedad (WHO, 2016a).

1.5.4.3 Análisis de sensibilidad:

El análisis de sensibilidad en el QMRA se utiliza para identificar factores de riesgo o exposición y ayuda en el desarrollo de prioridades para mitigar el riesgo. El análisis de sensibilidad juega un rol importante a la hora de verificar y validar el modelo, lo que proporciona más confianza a la hora de la toma de decisiones (Frey and Patil, 2002).

El objetivo del análisis de sensibilidad es evaluar el nivel de incertidumbre de cada parámetro introducido y como éste afecta el resultado final del análisis. Existen en la actualidad diferentes métodos matemáticos, estadísticos y gráficos disponibles para realizarlo. La selección del método

más adecuado dependerá de múltiples factores como el objetivo, el nivel de detalle requerido, el tipo de información utilizada en el modelo y la disponibilidad de programas informáticos (WHO, 2016a).

Un método bastante simple consiste en calcular la correlación de Spearman entre cada variable de entrada y el resultado del análisis de riesgo. Sus resultados se suelen representar través de un gráfico de tornado (Vose, 2008). El programa informático QMRAspot del Instituto Nacional de Salud Pública y Medio Ambiente de Holanda (RIVM), implementa otro método alternativo que consiste en evaluar la varianza de cada uno de los parámetros introducidos en cada paso. Todas las estimaciones obtenidas a través de la simulación de Monte Carlo son transformadas a su base logarítmica, se calculan las varianzas y estas son divididas por la varianza del indicador de salud establecido para caracterizar el riesgo (Schijven et al., 2014). Un método más sofisticado es el análisis de Sobol, el cual es considerado un análisis global, que consiste en descomponer la varianza del resultado final en fracciones asociadas a cada parámetro introducido en el modelo y a todas las posibles interacciones de cualquier orden entre ellos (Frey and Patil, 2002).

1.6 Técnicas de secuenciación de nueva generación aplicadas a la detección de virus

El veloz progreso en las tecnologías de secuenciación masiva asociadas con metodologías bioinformáticas han permitido una visión más detallada de la estructura y función de las comunidades virales, lo que ha favorecido la caracterización de virus emergentes (Ogilvie and Jones, 2015). Con el advenimiento de estudios de metagenómica, nuestro conocimiento de los diferentes componentes y la complejidad del microbioma se expanden exponencialmente (Hugenholtz and Tyson, 2008).

Las principales ventajas de la secuenciación masiva para caracterizar nuevos virus son la enorme cantidad de información de secuencias virales obtenidas y la habilidad de descubrir nuevos virus divergentes partiendo de secuencias previamente conocidas. Dentro de las desventajas podemos incluir la dificultad de separar el ácido nucleico viral del ácido nucleico del hospedador, por lo que es necesario el uso de filtros para eliminar residuos celulares y tratamientos de digestión enzimática para quedarnos con el material encapsidado. Finalmente, se tiene que incluir el coste que requiere la aplicación de esta técnica (Delwart, 2007).

1.6.1 Técnicas de secuenciación masiva del ADN

Las técnicas de secuenciación masiva son un conjunto de métodos que han surgido en los últimos años, colectivamente llamados como NGS por su nombre en inglés como "next-negeration sequencing" y particularmente conocidos como la segunda generación de métodos de secuenciación (la primera generación incluye al método de secuenciación sanger) (Pevsner, 2015). Los principales métodos de secuenciación masiva con sus principales características se describen en la Tabla 2.

Tabla 2: Comparación de los métodos de secuenciación masiva y de la secuenciación sanger (Pevsner, 2015).

Tecnología	Longitud de los reads (pb)	Reads por run	Tiempo del run	Coste por megabase (US$)	Exactitud
Roche 454	700	1 millón	1 día	10	99.9
Illumina	50 - 250	< 3 billones	1-10 días	0.10	98
SOLiD	50	1.4 billones	7-14 días	0.13	99.9
Ion Torren	200	< 5 millones	2 horas	1	98
Pacific Biosciences	2900	< 75000	< 2 horas	2	99
Sanger	400 - 900	N/A	< 3 horas	2400	99.9

1.6.1.1 Secuenciación con la plataforma de Ilumina

La plataforma de NGS más utilizada actualmente es probablemente Ilumina. Este método puede generar una terabase de datos de secuenciación de ADN en un simple run. Es un método muy fácil y ha permitido generar el 80% de toda la información obtenida a partir de los métodos de secuenciación masiva (Pevsner, 2015).

Para la preparación de la librería se requiere ADN fragmentado al que se le incorporan adaptadores específicos para indexar la muestra. Este proceso dependerá de la cantidad de ácido nucleico en la muestra y las opciones de sistemas disponibles más comunes son los protocolos Nextera, Nextera XT y TrueSeq. El sistema Nextera, utilizado en esta tesis, se basa en la utilización de transposones que por una reacción de "tagmentación" ponen o "etiquetan" el ADN de doble

cadena de la muestra con unos adaptadores a ambos extremos. Posteriormente, estos adaptadores servirán para realizar una reacción de PCR. En función de la cantidad de ADN disponible en la muestra se utiliza el kit Nextera (50ng ADN) o el Nextera XT (1ng) (Fernandez-Cassi, 2017).

El método de secuenciación masiva Ilumina trabaja bajo el principio de ciclos de terminación reversibles (Figura 1). Las hebras de ADN de la librería previamente indexada y con los adaptadores en ambos extremos son separadas y adheridas covalentemente a unos canales en una celda de flujo. A esta celda se le adicionan la ADN polimerasa y los desoxinucleótidos no marcados que crearán un "puente de amplificación", el cual une los dos extremos del ADN a la celda formando una U. Inmediatamente se genera una doble cadena que es desnaturalizada, de esta manera se generan "clusters" de cadenas simples que se encuentran densamente agrupadas en la celda. Una vez formados estos "clusters", se adicionan desoxinucleótidos marcados con un terminador. Al igual que en la secuenciación Sanger, este terminador no permitirá más elongación de la hebra, lo que permitirá que un láser excite las moléculas en cada "clúster" y se pueda leer la primera base que se adhirió. Finalmente el terminador es removido y el ciclo se vuelve a repetir (Pevsner, 2015).

Los servicios de Ilumina disponen de una gran diversidad de opciones de secuenciación que van en función del número y la longitud de los reads; así tenemos: Mini-Seq (8Gbp, $25x10^6$ reads, 2x150 paired-ends), Mi-Seq (15Gbp, $25x10^6$ reads, 2x300 paired-ends) y Hi-Seq (1500Gbp, $5x10^9$ reads, 2x150 paired-ends) (Fernandez-Cassi, 2017).

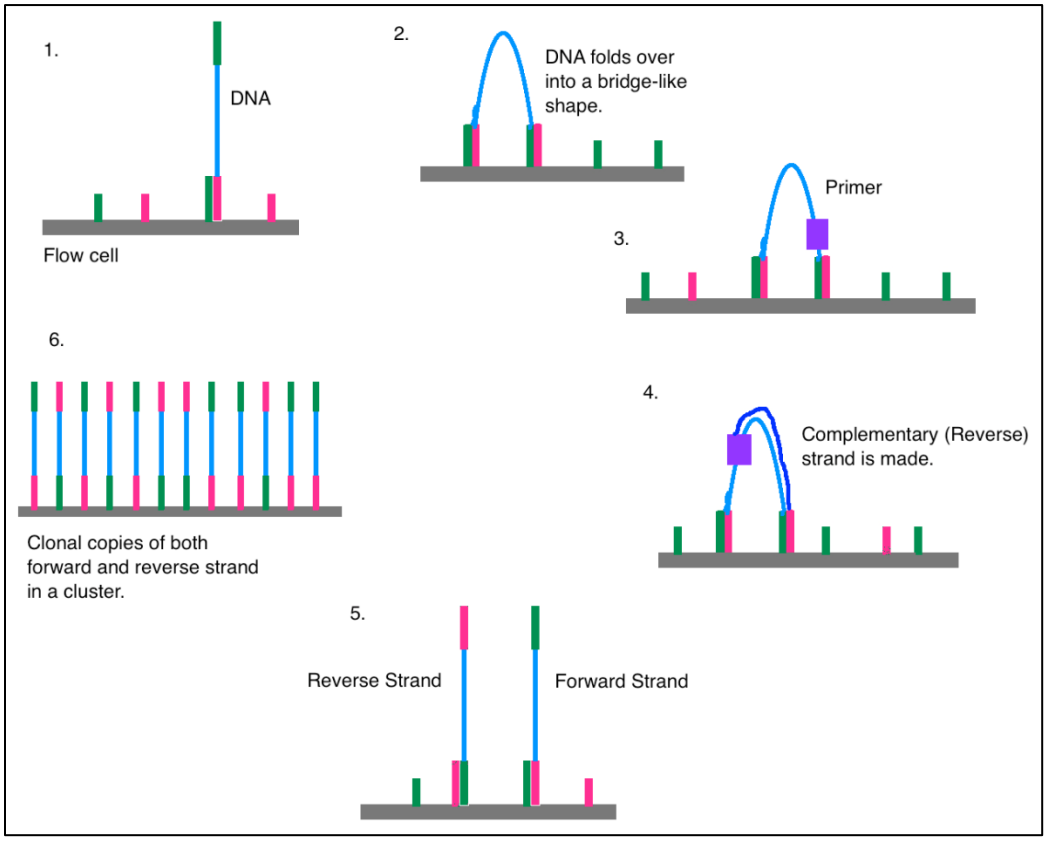
1.6.2 Bioinformática y análisis de los datos de la secuenciación masiva

Existen diversas herramientas informáticas -software local o servidores en línea- que permiten generar contigs a través del solapamiento de secuencias obtenidas con la secuenciación masiva. El uso de estas herramientas computacionales requiere de algoritmos de búsqueda como el BLAST para detectar similitudes con las secuencias virales conocidas, lo que puede requerir mucho tiempo. Además, es necesario definir los criterios para la clasificación de las secuencias dentro de los distintos grupos virales, uno de los más utilizados es el "E score" $<10^{-5}$, comúnmente usado para evaluar secuencias de origen viral (Delwart, 2007).

Un tema que suscita gran controversia es como definir que se ha detectado una secuencia viral nueva cuando se conoce una alta divergencia dentro de las bases de datos actuales. Se estima que

una gran fracción (entre 5% y 30%) de las secuencias obtenidas a partir de muestras de animales utilizando métodos de secuenciación masiva, no muestran ninguna relación significativa con las actuales secuencias presentes en el Genbank (Delwart, 2007).

Figura 1: Descripción gráfica de los ciclos que componen el proceso de amplificación de la técnica de secuenciación masiva Ilumina. Tomado de: DMLapato - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=43777596



El análisis de los datos obtenidos mediante la secuenciación masiva requiere un conjunto de pasos los cuales suelen representarse mediante un diagrama de flujo. Este diagrama describe una serie de pasos que se realizan con diferentes programas informáticos y se encargan de analizar las secuencias crudas, ensamblaje, alineamiento e interpretación de los datos. Algunos de los programas utilizados en el estudio de secuencias son FASTX-Toolkit, BWA, Bowtie, CLCbio, Metavelvet, Blast, Geneious, entre otros. Para más detalle se puede consultar la tesis de Natalia Timoneda, 2017 (Timoneda, 2017).

2 OBJETIVOS

Los objetivos planteados en esta tesis están orientados al estudio del riesgo asociado a la contaminación fecal del agua y alimentos. Este trabajo se divide en tres secciones: La primera parte busca mejorar las herramientas para cuantificar de una manera más precisa los patógenos y el riesgo microbiológico que tiene el consumo de agua contaminada. En la segunda parte se estima el riesgo microbiológico asociado a la utilización de agua regenerada para el riego de vegetales. Y en la última parte estudiamos pacientes con hepatitis aguda de etiología desconocida para la valoración de virus emergentes o virus nuevos de posible transmisión fecal-oral como posibles causantes de hepatitis.

Los objetivos específicos fueron:

- Determinar la eficiencia del SMF como método de concentración simultanea para virus, bacterias y protozoos.
- Comparar los métodos moleculares con los métodos de infectividad para estimar concentraciones virales en el agua.
- Evaluar un método de extrapolación para predecir intervalos que nos permitan corregir las cuantificaciones puntuales obtenidas a partir de métodos moleculares.
- Definir el nivel de variabilidad e incertidumbre del método de concentración con SMF para su aplicación a un estudio de análisis de QMRA.
- Modelizar la concentración viral de agua regenerada en función de la concentración del agua residual y la eficiencia del tratamiento para dos plantas de tratamiento y para dos virus diferentes (HAdV y NoV GII).
- Describir el riesgo de gastroenteritis al consumir vegetales regados con agua regenerada en función de la modelización de la concentración viral y la aplicación de un modelo desarrollado de QMRA.
- Identificar virus de posible transmisión fecal-oral como posibles causantes de hepatitis clínica aguda de etiología desconocida mediante la aplicación de técnicas de secuenciación masiva.

3 INFORMES

3.1 Informe de participación

El doctorando Eloy Anibal Gonzales Gustavson ha participado en los artículos que forman parte de su tesis doctoral de la manera que se detalla a continuación:

Gonzales-Gustavson, Eloy; Cárdenas-Youngs, Y; Calvo, M; da Silva, M; Hundesa, A; Amorós, I; Moreno, Y; Moreno-Mesonero, L; Rosell, R; Ganges, L; Araujo, R; Girones, R. Characterization of the efficiency and uncertainty of skimmed milk flocculation for the simultaneous concentration and quantification of water-borne viruses, bacteria and protozoa. Journal of Microbiological Methods (2017) 134:46-53. DOI: 10.1016/j.mimet.2017.01.006.

El doctorando llevó a cabo la coordinación y el desarrollo del trabajo experimental y procesamiento de las muestras virológicas excepto la parte del cultivo del BVDV. También estuvo involucrado en el análisis estadístico de los datos y en la elaboración de las tablas y parte de las figuras. Finalmente se encargó de la redacción del artículo y su presentación a la revista bajo la tutela de los directores.

Gonzales Gustavson, E.; Calvo, M; Rusiñol M; Medema G; R. Girones. Quantitative risk assessment for the use of reclaimed water to irrigate lettuce in Catalonia. (En preparación).

El doctorando llevó a cabo el análisis completo del QMRA, análisis estadístico de los datos con la dirección de los directores de tesis y en la elaboración de las tablas y figuras. Finalmente se encargó de la redacción del artículo y su presentación a la revista bajo la tutela de los directores.

Gonzales Gustavson, E., N. Timoneda, X. Fernandez-Cassi, A. Caballero, J. F. Abril, M. Buti, F. Rodriguez-Frias, R. Girones. Identification of sapovirus GV.2, astrovirus VA3 and novel

anelloviruses in serum from patients with acute hepatitis of unknown aetiology. PLoS ONE 2017 Oct 5;12(10):e0185911. DOI: 10.1371/journal.pone.0185911

El doctorando llevó a cabo parte del trabajo experimental y procesamiento de muestras. También se encargó del análisis de las secuencias, interpretación de los datos y elaboración de los arboles filogenéticos. Finalmente se encargó de la redacción del artículo y su presentación a la revista, con la dirección de los directores de tesis.

Rosina Girones Llop Miquel Calvo

Barcelona, 27 de noviembre del 2017

3.2 Informe sobre el factor de impacto de las publicaciones

Los artículos que forman parte de la memoria de la tesis doctoral presentada por Eloy Anibal Gonzales Gustavson han estado publicados o sometidos para su publicación en revistas internacionales indexadas tal como se detalla a continuación:

El artículo: "Characterization of the efficiency and uncertainty of skimmed milk flocculation for the simultaneous concentration and quantification of water-borne viruses, bacteria and protozoa" se publicó en la revista *Journal of Microbiological Methods* el año 2017 en formato OPEN ACCESS y con un índice de impacto de 2,09.

El artículo: "Identification of sapovirus GV.2, astrovirus VA3 and novel anelloviruses in serum from patients with acute hepatitis of unknown aetiology" se publicó en la revista *PLOS ONE* el año 2017 en formato OPEN ACCESS y con un índice de impacto de 3,54.

El artículo: "Characterizing the concentration of norovirus and adenovirus in reclaimed water and assessment of the risk from their use to irrigation of lettuce in Catalonia" está sometido para su publicación.

Rosina Girones Llop Miquel Calvo

Barcelona, 27 de noviembre del 2017

4 PUBLICACIONES

4.1 Artículo 1: Caracterización de la eficiencia e incertidumbre de la floculación con leche desnatada para la concentración simultánea y cuantificación de virus, bacterias y protozoos

Characterization of the efficiency and uncertainty of skimmed milk flocculation for the simultaneous concentration and quantification of water-borne viruses, bacteria and protozoa.

Gonzales-Gustavson, Eloy; Cárdenas-Youngs, Y; Calvo, M; da Silva, M; Hundesa, A; Amorós, I; Moreno, Y; Moreno-Mesonero, L; Rosell, R; Ganges, L; Araujo, R; Girones, R.

En este estudio se evaluó el uso del método de concentración de floculación orgánica con leche descremada para concentrar simultáneamente virus, bacterias y protozoos. Para este fin, se seleccionaron bacterias indicadoras de contaminación fecal y patógenos como *E. coli* y *H. pylori*, los virus HAdV 35, RoV SA-11, el bacteriófago MS2 y el virus con envoltura BVDV; y los protozoos seleccionados fueron *Acanthamoeba, Giardia* y *Criptosporidium*. Estos microorganismos fueron cuantificados e inoculados en muestras de 10 litros de agua potable previamente tratadas con tiosulfato de sodio. Las muestras se concentraron mediante el método de SMF y se volvieron a cuantificar los microorganismos en el concentrado; en el caso de los virus se realizó doble cuantificación: por q(RT)PCR y a través de un método de cuantificación de virus infecciosos en líneas celulares. Los porcentajes medios de recuperación obtenidos mediante q(RT)PCR fueron 66% (HAdV 35), 24% (MS2), 28% (RoV SA-11), 15% (BVDV), 60% (*E. coli*), 30% (*H. pylori*) y 21% (*A. castellanii*). Los porcentajes de recuperación obtenidos mediante infectividad fueron: 59% (HAdV 35), 12% (MS2), 26% (RoV SA-11) y 0.7% (BVDV). Los porcentajes de recuperación para los protozoos *Giardia* y *Criptosporidium*, que se midieron a través de la técnica de inmunofluoroscencia, fueron 18% y 13%, respectivamente. Aunque en general las cuantificaciones fueron más altas mediante el método de cuantificación de q(RT)PCR, los porcentajes de recuperación fueron similares para HAdV 35 y para RoV SA-11. Además, se modelizó la variabilidad y la incertidumbre de los valores de recuperación obtenidos para aplicarlos a un método de

extrapolación de las cuantificaciones obtenidas mediante q(RT)PCR y poder obtener la concentración real. Se obtuvieron intervalos de predicción al 95% de la concentración real para los microorganismos evaluados por q(RT)PCR mediante el método general de “bootstrap” no paramétrico que fue adaptado en nuestro contexto para estimar el error técnico de las mediciones. El método de concentración SMF mostró porcentajes de recuperación con baja variabilidad que permitió el uso de una aproximación matemática para predecir la concentración real de patógenos e indicadores con intervalos aceptablemente bajos. Estas estimaciones son de utilidad en estudios de QMRA y en el control de la calidad microbiológica del agua.

Journal of Microbiological Methods 134 (2017) 46–53

Contents lists available at ScienceDirect

Journal of Microbiological Methods

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

Characterization of the efficiency and uncertainty of skimmed milk flocculation for the simultaneous concentration and quantification of water-borne viruses, bacteria and protozoa

Eloy Gonzales-Gustavson [a], Yexenia Cárdenas-Youngs [a], Miquel Calvo [a], Marcelle Figueira Marques da Silva [a], Ayalkibet Hundesa [a], Inmaculada Amorós [b], Yolanda Moreno [b], Laura Moreno-Mesonero [b], Rosa Rosell [c], Llilianne Ganges [c], Rosa Araujo [a], Rosina Girones [a,*]

[a] Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, Av. Diagonal 643, 08028 Barcelona, Catalonia, Spain
[b] Instituto de Ingeniería del Agua y Medio Ambiente, Ciudad Politécnica de la Innovación, Ed. 8G, Acceso D, planta 2, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain
[c] IRTA, Centre de Recerca en Sanitat Animal (CReSA, IRTA-UAB), Campus de la Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

ARTICLE INFO

Article history:
Received 11 October 2016
Received in revised form 5 January 2017
Accepted 8 January 2017
Available online 16 January 2017

Keywords:
Skimmed milk flocculation
Water contamination
Microorganisms
q(RT)PCR
Recovery
Real concentration

ABSTRACT

In this study, the use of skimmed milk flocculation (SMF) to simultaneously concentrate viruses, bacteria and protozoa was evaluated. We selected strains of faecal indicator bacteria and pathogens, such as *Escherichia coli* and *Helicobacter pylori*. The viruses selected were adenovirus (HAdV 35), rotavirus (RoV SA-11), the bacteriophage MS2 and bovine viral diarrhoea virus (BVDV). The protozoa tested were *Acanthamoeba*, *Giardia* and *Cryptosporidium*. The mean recoveries with q(RT)PCR were 66% (HAdV 35), 24% (MS2), 28% (RoV SA-11), 15% (BVDV), 60% (*E. coli*), 30% (*H. pylori*) and 21% (*Acanthamoeba castellanii*). When testing the infectivity, the mean recoveries were 59% (HAdV 35), 12% (MS2), 26% (RoV SA-11) and 0.7% (BVDV). The protozoa *Giardia lamblia* and *Cryptosporidium parvum* were studied by immunofluorescence with recoveries of 18% and 13%, respectively. Although q(RT)PCR consistently showed higher quantification values (as expected), q(RT)PCR and the infectivity assays showed similar recoveries for HAdV 35 and RoV SA-11. Additionally, we investigated modelling the variability and uncertainty of the recovery with this method to extrapolate the quantification obtained by q(RT)PCR and estimate the real concentration. The 95% prediction intervals of the real concentration of the microorganisms inoculated were calculated using a general non-parametric bootstrap procedure adapted in our context to estimate the technical error of the measurements. SMF shows recoveries with a low variability that permits the use of a mathematical approximation to predict the concentration of the pathogen and indicator with acceptable low intervals. The values of uncertainty may be used for a quantitative microbial risk analysis or diagnostic purposes.

1. Introduction

Diseases related to water contamination constitute a major human health issue. Inadequate drinking water and poor sanitation are estimated to cause 842,000 diarrhoeal disease-related deaths per year (World Health Organization, 2014). They are related to a broad range of health problems and cause impacts on productivity due to water-borne diseases (Amini and Kraatz, 2014). Moreover, the creation of protocols to measure water quality, considering the diversity of pathogens that may be present, is one of the major problems that must be solved for improving the control of water quality and Quantitative Microbial Risk Assessment (QMRA) studies.

The following four main critical steps in the process of evaluating the microbiological quality of water need to be considered: (1) which pathogens may be present; (2) which microorganisms are used as indicators of contamination; (3) which method is used to concentrate the particular indicator or indicators; and (4) which technique is used to detect them.

Indicator organisms are used for a range of purposes as follows: indicators of faecal pollution and to evaluate the effectiveness of processes such as filtration or disinfection. The most popular indicator organisms are thermotolerant coliforms, *E. coli* and intestinal enterococci. However, the suitability of *E. coli* as an indicator has been questioned, because its survival in water and sensitivity to treatment and disinfection processes differ substantially from those of excreted viruses and protozoa.

Abbreviations: BVDV, Bovine viral diarrhoea virus; HAdV, Human adenovirus; IFA, Immunofluorescence assays; JCPyV, JC polyomavirus; NoV, Norovirus; PI, Prediction intervals; QMRA, Quantitative microbial risk assessment; q(RT)PCR, Quantitative (reverse transcriptase) PCR; RoV, Rotavirus; SMF, Skimmed milk flocculation; TCID50, 50% Tissue culture infective dose.

* Corresponding author.
E-mail address: rgirones@ub.edu (R. Girones).

http://dx.doi.org/10.1016/j.mimet.2017.01.006

E. coli is far more sensitive, and the consequence is a low correlation with the presence of other pathogens (Amini and Kraatz, 2014; Bofill-Mas et al., 2013; Gorchev and Ozolins, 2011).

Coliphages share many properties with human viruses and are used as models to assess the behaviour of excreted viruses in the water environment. In this regard, they are superior to faecal bacteria. However, there is no direct correlation between the numbers of coliphages and the numbers of excreted viruses (Gorchev and Ozolins, 2011). The use of excreted viruses as microbial indicators is based on the shortcomings of the existing choices. Human adenovirus (HAdV) has been proposed as a viral indicator of contamination (Gorchev and Ozolins, 2011; Pina et al., 1998) and has been used in various studies as a viral indicator of human faecal contamination and a microbial source tracking tool (Bofill-Mas et al., 2011, 2013; Rusiñol et al., 2014).

Most HAdVs are associated with respiratory disease, but types 40 and 41 are responsible for gastroenteritis outbreaks in children (Wold and Horwitz, 2013). Rotavirus (RoV) is also associated with gastroenteritis; RoV-A is the most common cause of severe vomiting and diarrhoea among children up to 30 months old (Estes and Greenberg, 2013). The coliphage MS2 is commonly used as a surrogate and process control in microbiological food and water analyses (van Duin and Olsthoorn, 2012). BVDV is an important cause of morbidity, mortality, and economic loss in dairy and beef cattle worldwide (MacLachlan and Dubovi, 2011).

E. coli is commonly found in the lower intestine of warm-blooded organisms. Most strains are harmless, but others can cause serious food poisoning and are responsible for product recalls due to food contamination (Madigan et al., 2014). *H. pylori* is an acid-tolerant bacterium usually found in the stomach and is related to gastric cancer (Johnson et al., 1997). *H. pylori* has been detected in wastewater (Moreno and Ferrús, 2012), surface water and other environmental samples all over the world (Eusebi et al., 2014) and has even demonstrated the capacity to survive in chlorinated water when the enumeration of coliforms indicates that the water is potable (Santiago et al., 2015). *G. lamblia* and *C. parvum* are responsible for outbreaks of gastroenteritis related to the consumption of contaminated water (Gascón, 2006). *Acanthamoeba* spp., free-living protozoa, are considered to be opportunistic pathogens (Marciano-Cabral and Cabral, 2003) and are known to have a role in the persistence of some bacterial pathogens, such as *Legionella*, in water environments (Lambrecht et al., 2015).

The direct examination of water is difficult due to low and fluctuating concentrations of microorganisms and because concentration procedures are usually organism and/or matrix-specific and most techniques have high or unknown variability parameters. One-step skimmed milk flocculation (SMF) has been proposed as an efficient low-cost method to concentrate viruses in all types of water samples. This method has been used in environmental water matrices such as river water (Calgua et al., 2013a), seawater (Calgua et al., 2008), ground water (Bofill-Mas et al., 2011) and wastewater (Calgua et al., 2013b). However, the efficacy of the recovery in controlled conditions has not been properly described until now.

Quantitative Microbial Risk Analysis (QMRA) is a scientific tool used to assess the microbial safety of water and is needed for developing a strategy of risk management models. QMRA models each variable using a probability distribution. The advantage is that the result is represented by a probability distribution function instead of a single value. The objective of QMRA is the ability to calculate the combined impact of the uncertainty in the model's parameters to determine an uncertainty distribution of the possible model outcomes (Vose, 2008).

The aim of the present study was to determine the efficacy of the SMF recovery to simultaneously concentrate viruses, bacteria and protozoa and then compare q(RT)PCR and infectivity assays to detect and quantify the number of viruses recovered. Finally, an extrapolation method was evaluated with the q(RT)PCR quantification using the prediction interval (PI) based on the known recoveries to correctly achieve the actual concentration of the spiked microorganisms and define the uncertainty values of the method.

2. Materials and methods

2.1. Microorganism stocks and cell lines

The following viruses were analysed and spiked into the water samples: HAdV-35 (ATCC, LGC Standards AB, Borås, Sweden) cultured in cell line A549 (ATCC CCL-185), MS2 (ATCC 23631) cultured in *Salmonella typhimurium* strain WG49 (NCTC 12484), RoV SA-11 (ATCC VR-1565) cultured in MA104 (ATCC CRL-2378) and Bovine viral diarrhoea virus (BVDV) strain NADL kindly donated by the EU and OIE Reference Laboratory for Classical Swine Fever, Institute of Virology, University of Veterinary Medicine, Hannover, Germany, and cultured in NDBK (ATCC CCL-22). The analysed bacteria were *E. coli* (ATCC 23725) and *H. pylori* (NCTC11637). The protozoa tested in the study were *A. castellanii* (CCAP 1534/2), *G. lamblia* H3 isolate (Waterborne Inc., New Orleans, LA) and a *C. parvum* Iowa isolate (Waterborne Inc., New Orleans, LA).

2.2. Water samples

This experiment was conducted with tap water from the metropolitan area of Barcelona; the volume of water evaluated in each bucket was 10 L. The number of buckets inoculated with each of the microorganisms and their respective inoculated concentration are specified in Table 1. The tap water was previously treated with 100 mL of sodium thiosulfate (10% (*w*/*v*)) to eliminate chloride residues. Four additional buckets with the same volume of water were analysed as negative control samples.

2.3. Skimmed milk flocculation concentration

The skimmed milk flocculation concentration protocol has been described in previous studies (Bofill-Mas et al., 2011; Calgua et al., 2008). In summary, a pre-flocculated skimmed milk solution (1% (*w*/*v*)) was prepared by dissolving 10 g of skimmed milk powder (Difco-France) in 1 L of artificial seawater and carefully adjusting the pH to 3.5 with 1 N HCl. One hundred millilitres of this solution was added to each of the previously acidified (pH 3.5) 10 L water samples (the final concentration of skimmed milk was 0.01% (*w*/*v*)). The conductivity was also measured and adjusted with artificial sea salt (Sigma, Aldrich Chemie GMBH, Steinheim, Germany) to achieve a minimum conductivity of 1.5 mS/cm^2. The samples were stirred for 8 h at room temperature, and the flocs were allowed to settle by gravity for another 8 h. The supernatants were removed, and the sediment was collected and transferred to a 500 mL centrifuge container and centrifuged at 8000 × *g* for 30 min at 4 °C. The obtained pellet was resuspended in 8 mL of

Table 1
The number of microorganisms inoculated in each of the ten litre water buckets used for the skimmed milk flocculation concentration process.

Microorganisms (number of samples)	Molecular quantification		Quantification by infectious assays	
HAdV (10)	2.88E + 07	GC	4.60E + 06	IFA
MS2 (13)	2.92E + 09	GC	2.07E + 09	PFU
	2.92E + 07	GC	5.03E + 06	
RoV (19)	6.31E + 08	GC		
	2.09E + 07	GC	4.08E + 05	TCID50
BVDV (3)	2.10E + 08	GC	6.31E + 05	TCID50
E. coli (10)	2.37E + 06	GC		
H. pylori (9)	1.97E + 08	GC		
A. castellanii (9)	7.27E + 04	GC		
C. parvum (8)	1.46E + 04	IFA		
G. lamblia (8)	1.56E + 04	IFA		

GC: genomic copies; IFA: immunofluorescence assay; PFU: plaque-forming units; TCID50: 50% tissue culture infective dose.

0.2 M phosphate buffer at pH 7.5 (1:2, *v*/*v* of 0.2 M Na_2HPO_4 and 0.2 M NaH_2PO_4). Once the pellet was completely dissolved, the phosphate buffer was added to a final volume of 10 mL. The concentrates were kept at −20 °C after the SMF method was performed. The quantification was then performed within five days.

2.4. Nucleic acid extraction

Viral nucleic acids (NA) were extracted using the QIAmp Viral RNA kit (Qiagen, Inc., Valencia, CA). Bacterial and protozoan DNA was extracted using the UNEX method (Hill et al., 2015). The volumes of the concentrates used for the extraction were 140 and 300 μL, and the elutions were 80 and 100 μL, for viruses/bacteria and protozoa, respectively. Immediately after extraction, q(RT)PCR analyses were performed.

2.5. q(RT)PCR quantification

Specific real-time q(RT)PCR assays were used to quantify the microorganisms following the specifications previously described for HAdV (Hernroth et al., 2002), RoV (Zeng et al., 2008), MS2 (Calgua et al., 2014), BVDV (Losurdo et al., 2015), *E. coli* (Khan et al., 2007), *H. pylori* (Santiago et al., 2015) and *A. castellanii* (Qvarnstrom et al., 2006). Undiluted and 10-fold dilutions of the nucleic acid extracts were analysed in duplicate, including the concentrates from negative control buckets. All of the q(RT)PCR assays included four non-template controls to demonstrate that the mix did not produce fluorescence. The standards for viruses were prepared using synthetic gBlocks® Gene Fragments (IDT) and quantified with a Qubit® fluorometer (Thermo Fisher Scientific). For bacteria and *A. castellanii* standards, the DNA was extracted from cultures of known concentration and quantified using a Nanodrop 1000. For all of the standards, ten-fold dilutions were prepared from 10^0 to 10^7 copies per reaction.

2.6. Infectivity and immunofluorescence assays

Specific infectivity assays were performed using previously described methods for viruses as follows: IFA for HAdV-35 (Calgua et al., 2011), TCID50 for RoV (Otto et al., 2015) and BVDV (OIE, 2015) and plaque assays for MS2 (Anonymous, 1995).

For the quantification of *G. lamblia* and *C. parvum*, we used an immunofluorescence method previously described for the staining of cysts and oocysts with 4′,6-diamidino-2-phenylindole (DAPI) and observed the staining using differential interference contrast microscopy (USEPA, 2005).

2.7. Recovery, concentration and PI estimates

In our assays, in every replicate, the true concentration, Y_K, was known, allowing us to compute the recovery mean $\overline{r_K}$. In a future non-controlled experiment, the only information available will be the q(RT)PCR-measured concentration, m_U, while the concentration Y_U and the recovery r_U will be unknown. In this section, we introduce a new PI, which estimates this unknown concentration Y_U of new observations, guaranteeing an $1-\alpha$% confidence level.

Our approach starts by considering the distribution of our controlled recoveries as a valid model for the future measures. In a new experiment (with one replicate), the three quantities are related by the equation $m_U = Y_U r_U$. A *point estimate* of Y_U can be obtained by simply substituting r_U with $\overline{r_K}$:

$$\hat{Y}_U = \frac{m_U}{\overline{r_K}} \tag{1}$$

The *relative error* of this estimation is:

$$e_R = \frac{\hat{Y}_U}{Y_U} = \frac{r_U}{\overline{r_K}} \tag{2}$$

We consider it essential to improve Eqs. (1) and (2) by also measuring their confidence: the accuracy of $\overline{r_K}$ depends on the sample size and variability of our current experiments.

The PIs described in the statistical literature are built specifically to predict new observations when the parameters of the distribution are estimated with a sample. The purpose and formulas of the PIs are different from the more commonly used confidence intervals. In the Gaussian case, the PI has a closed simple expression. However, the normality assumption for the recoveries is not supported in practice, and other probabilistic models are often used; for instance, Petterson et al. (2015) assumed the beta distribution to study the variability in the recovery of a virus in water.

The absence of closed expressions for the PI plus the difficulty in ensuring a correct goodness of fit of any probabilistic model has finally impelled us to find a free-distribution method. Among the different approaches previously described (see Bai et al., 1990; Mojirsheibani, 1998, for a comparison of several methods), we chose the non-parametric bootstrap-t technique. In brief, this standard computational method defines a bootstrap statistic T^*, which combines the distributions of the past and the future samples. Given a confidence level of $1-\alpha$%, the resampling procedure lets us obtain any α quantile $\hat{t}^{(\alpha)}$ of T^*, and, in our context, to obtain the following limits of the PI for r_U (further details in Mojirsheibani, 1998):

$$prob(r_U \le r_{U,\,min}) = \frac{\alpha}{2} \quad prob(r_U \le r_{U,\,max}) = 1-\frac{\alpha}{2} \tag{3}$$

The above bootstrap-t PI of r_U lets us derive from Eqs. (1) and (2) two new expressions: the PI of the unknown concentration and its relative error

$$prob\left(\frac{m_U}{r_{U,\,max}} \le Y_U \le \frac{m_U}{r_{U,\,min}}\right) = 1-\alpha \tag{4}$$

$$prob\left(\frac{r_{U,\,min}}{\overline{r_K}} \le e_R \le \frac{r_{U,\,max}}{\overline{r_K}}\right) = 1-\alpha \tag{5}$$

2.8. Statistical evaluation

All of the data were statistically analysed with the 3.1.1 version of the R software (R Core Team, 2016). First, descriptive statistics of all of the recoveries and quantifications were performed. We plotted the actual data together with the normal density, the beta density (estimated using the maximum likelihood criteria) and a non-parametric kernel estimation of the density in order to assess their goodness of fit. The PI (3), (4) and (5) for every organism were computed implementing the equations for T^* and $\hat{t}^{(\alpha)}$ combined with the methods of the *boot* package in R. An R script with our implementation of these PIs can be found in the supplementary material of this paper. Additionally, Wilcoxon signed-rank tests were used to evaluate the difference between the methods of quantification (q(RT)PCR and the infectivity assay) with the recoveries and quantification of RoV. Spearman's rank correlation coefficient was used to evaluate the q(RT)PCR recovery between all of the microorganisms evaluated.

3. Results

3.1. Recovery efficacy

Each water bucket was inoculated with the concentration of microorganisms indicated in Table 1. Bacteria and protozoa were quantified using one method (q(RT)PCR or the infectivity/immunofluorescence assays), whereas viruses were quantified with both methods. The recovery percentage for each microorganism represents the efficacy of recovery using the combination of SMF plus the efficacy of the method of quantification, either q(RT)PCR or the infectivity assay (Table 2). All of the negative control buckets were negative.

3.2. Correlation between the recoveries for the different microorganisms

The recovery results of each of the samples were correlated to determine if there are relationships between the microorganisms for the q(RT)PCR results. Table 3 shows the results of the Spearman's correlation analysis of the values obtained, specifying the number of samples and the *p*-values. It also includes a graphical representation of the data in Cartesian planes inside a correlation matrix between each of the pairs compared. A positive correlation was found between the tested bacteria *H. pylori* and *E. coli* (Table 3). A positive but non-statistically significant correlation also occurred between all the viruses evaluated. *Acanthamoeba* recovery was not correlated with any microorganism evaluated by q(RT)PCR.

3.3. Comparison of the quantification between q(RT)PCR and the infectivity assays

Viruses were enumerated with both q(RT)PCR and infectivity assays; the results for the recoveries are shown in Table 2 and the quantifications in Table 4. The descriptive results of both tables must be carefully interpreted because the small sample size of the infectivity assays does not allow an inferential assessment. For RoV, where 10 replicates are available, the Wilcoxon signed-rank test was computed.

The recoveries with q(RT)PCR and the infectivity assays for RoV show non-significant differences (p-value = 0.37). The descriptive results of HAdV in Table 2 may suggest a similar conclusion. In contrast, also in RoV, we detected significant differences (p-value = 0.002) between the quantities obtained by q(RT)PCR and by infectivity. Table 4 may suggest similar results for the rest of the organisms, but further experiments are required to confirm these preliminary results.

A ratio between the logarithm of the quantification between q(RT)PCR and infectivity was calculated to indicate how many times the quantifications varied relative to one another. The ratios for HAdV, RoV and MS2 were 1.13, 1.35 and 1.07, respectively.

Table 2
Skimmed milk flocculation recoveries for each of the microorganisms evaluated.

Microorganisms	Method	Percent recovery					
		Mean %	CI 95% of mean	*n*	sd	min	max
HAdV	qPCR	66	53.5–78.5	10	17.4	32.2	86.7
	IFA	58.7	4.5–100	3	1.8	8.1	49.8
MS2	q(RT)PCR	23.9	19.6–28.1	13	7	13.8	36.8
	PFU	11.9	9–14.7	4	1.8	9.5	13.9
RoV	q(RT)PCR	28.2	25.6–30.7	19	5.3	16	37.1
	TCID50	26.1	17.1–35.1	10	12.6	43.5	83.7
BVDV	q(RT)PCR	14.7	10.8–18.7	3	1.6	12.9	15.8
	TCID50	0.7	0.4–1.1	3	0.13	0.67	0.89
E. coli	qPCR	59.6	40.3–79	10	27.1	15.6	98.7
H. pylori	qPCR	30.2	24.4–36.1	9	7.6	20.8	41.5
A. castellanii	qPCR	20.5	14.9–26.1	9	7.2	13	32.1
G. lamblia	IFA	17.8	15–20.7	8	3.4	12.8	21.5
C. parvum	IFA	12.8	12.5–15.2	8	2.9	9.6	17.4

q(RT)PCR: quantitative (reverse transcriptase) PCR; IFA: immunofluorescence assay; PFU: plaque-forming units; TCID50: 50% tissue culture infective dose.

3.4. The impact of recovery on the predicted concentration

The use of q(RT)PCR quantification in QMRA has been previously demonstrated (Rames et al., 2016). To better evaluate the real concentration of microorganisms when the quantification is obtained after SMF, we suggest extrapolation of the q(RT)PCR value incorporating the uncertainty and variability of the method. HAdV, RoV, MS2, *E. coli*, *H. pylori* and *A. castellanii* were used for this purpose. The upper and lower limits, including the real concentration in water samples with a 95% PI, were estimated using the non-parametric bootstrap approach described above. Despite the moderate sample size of our assays (between 9 and 19 replicates), the PIs show a reasonable width of approximately 4–5 units, supporting the applicability of this information for future observations.

As an example of how to use this information in practice, we take here the measurement previously published by Calgua et al. (2008) as a *future* measurement; their reported m_U was 2.73E + 4 genomic copies in 10 L of HAdV. Substituting in Eqs. (4) and (5) the values in Table 5 (bootstrap-t PI method, row HAdV) we obtain:

$$prob(2.73E+4\times 1.027 \le Y_U \le 2.73E+4\times 5.200) =$$

$$prob(2.80E+4 \le Y_U \le 1.42E+5) = 0.95$$

$$prob(29.14\% \le e_R \le 147.57\%) = 0.95$$

In fact, Calgua et al. (2008) state that the real concentration inoculated in the sample was 4.04E + 04 with 68% recovery. Both quantities lie in their respective PIs computed above.

In Table 5, we have additionally computed the PI when a normal distribution of the recoveries is assumed. In some organisms, this PI may show comparable results to the bootstrap-t PI; for instance, in RoV (Table 5) they are almost identical. Fig. 1A can explain this concordance: normal, beta and kernel densities similarly fit the actual data. This is not the general case and, as a consequence, the normal and bootstrap PIs may show different coverages. For instance, in HAdV, the different results in Table 5 can be explained by the different fittings of the densities in Fig. 1B. The main reason to introduce the bootstrap-t predictions was the lack of fit of the normal distribution needed to correctly build prediction intervals.

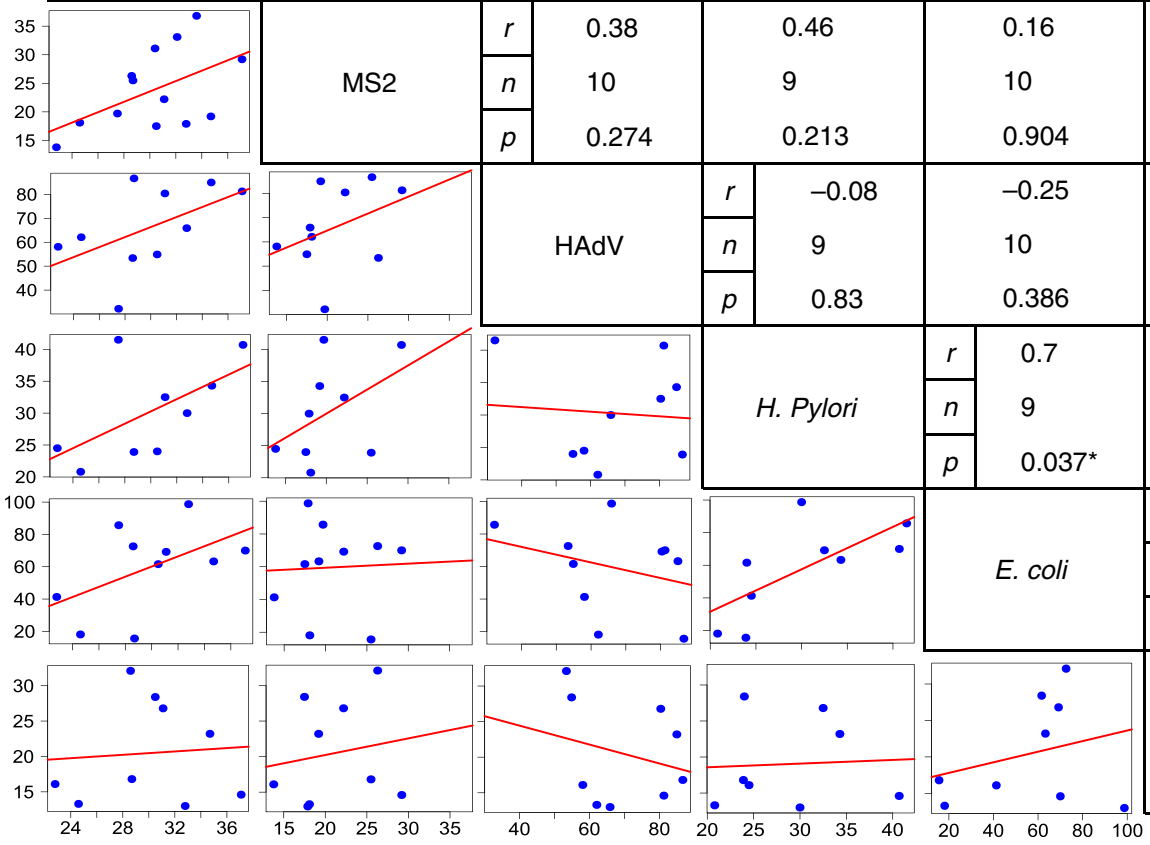
4. Discussion

This is the first study to characterize the simultaneous concentration of viruses (including an enveloped virus), bacteria and protozoa with SMF. Moreover, this is the first study to correlate recoveries and evaluate the uncertainty of the results when using SMF for simultaneous concentration. Additionally, the evaluation of infectious viral particles has been included, since there was no previous information of recoveries with SMF using infectivity assays. The highest recoveries obtained in this experiment were with HAdV and *E. coli*. This method has been previously evaluated by spiking HAdV in seawater, river water and wastewater with recoveries by qPCR of 52%, between 41% and 50%, and between 30% and 95%, respectively (Calgua et al., 2008, 2013a,b). Moreover, SMF has been used to concentrate norovirus (NoV) with recoveries between 34 and 74% (Calgua et al., 2013a). The present study describes the recoveries of HAdV with confidence intervals that include these previous results, with higher sample size and suggesting that SMF with qPCR quantification may be used indistinctly in different water matrices without affecting the efficacy of the method.

SMF has been used in environmental samples to detect DNA viruses used as microbial source tracking tools, such as JC polyomavirus (JCPyV), porcine adenovirus and bovine polyomavirus, in superficial and ground water samples (Bofill-Mas et al., 2011); it has also been used in studies analysing a wide diversity of viruses, such as HAdV, NoV, JCPyV, RoV, Klassevirus, Asfarvirus-like virus and Merckel cell

Table 3
Correlation of the recoveries obtained by q(RT)PCR between microorganisms with the skimmed milk flocculation method.

RoV	*r*	0.51	0.53	0.59	0.51	0.08
	n	13	10	9	10	9
	p	0.073	0.118	0.096	0.134	0.848
MS2	*r*	0.38	0.46	0.16	0.16	
	n	10	9	10	9	
	p	0.274	0.213	0.904	0.676	
HAdV	*r*	−0.08	−0.25	−0.25		
	n	9	10	9		
	p	0.83	0.386	0.52		
H. Pylori	*r*	0.7	0.06			
	n	9	8			
	p	0.037*	0.896			
E. coli	*r*	0.27				
	n	9				
	p	0.48				
A. castellanii						



r = Spearman's rank correlation coefficient, *n* = paired sample size, *p* = *p*-value, *p-value < 0.05.

polyomavirus in river water (Calgua et al., 2013a; Rusiñol et al., 2014, 2015) and HAdV, RoV, PP7 phage and NoV in seawater (Calgua et al., 2008; Rusiñol et al., 2014; Victoria et al., 2014). The SMF protocol with modifications has also been used to quantify HAdV, JCPyV and NoV in sewage water (Calgua et al., 2013b). Additionally, SMF has been modified to detect HAdV and NoV in strawberries with good results (Melgaço et al., 2016).

Enveloped viruses such as BVDV may be more stable than expected in water. Considering the lack of information available on the concentration protocols of enveloped viruses in water, it was decided to include in this study a representative enveloped virus, BVDV, an important pathogen for cattle. The recoveries of this virus using SMF were analysed in triplicate, and the applicability of available methodologies, specifically qPCR and infectivity assays, was also evaluated. The analysis of BVDV, transmitted through inhalation and ingestion as main horizontal routes in cattle (MacLachlan and Dubovi, 2011), in water will produce useful information on the spread of BVDV through contaminated sources of water and animal drinking troughs.

Table 4
A comparison between q(RT)PCR and infectivity results in the viral concentrates after skimmed milk flocculation.

Virus	*N*	Quantification		
		q(RT)PCR	Infectivity	Log 10 ratio q(RT)PCR/infectivity
RoV	10	6.25E + 06	1.06E + 05	1.35
MS2	4	7.38E + 08	1.83E + 08	1.07
HAdV	3	1.87E + 07	2.70E + 06	1.13
BVDV	3	3.09E + 07	4.67E + 03	2.04

The availability of cost-effective techniques for the simultaneous concentration of viruses, bacteria and parasites from water will be very useful when the monitoring of microbial water quality for diverse microbe types is desired (Hill et al., 2005). In addition, it will also be valuable for the application of next-generation sequencing methods and the characterization of the microbial population of water. Until now, there have been no other methods of concentration that allow for the evaluation of a representative volume (10 L), diverse water matrices with high and low turbidity, a high recovery percentage and the simultaneous evaluation of viruses, bacteria and protozoa.

Due to the importance of finding a suitable indicator of contamination, the correlation between the recoveries of the different pathogens and suggested indicators is relevant information that must be evaluated. In theory, flocs adsorb particles in an acid medium, which increases their weight and facilitates the precipitation of the particles over time (Calgua et al., 2008). These results suggest that the efficacy of the flocs to aggregate the particles in a water suspension may show small changes depending on the type of microorganism. The results also suggest that a single faecal indicator is not feasible; however, the correlations

Table 5
Values used to compute a prediction interval for Y_u with a 95% confidence level (column 2.5% shows $(\frac{1}{r_{U,max}})$, 97.5% $(\frac{1}{r_{U,min}})$).

Virus	n	Bootstrap-t		Normal (unknowns μ and σ)	
		2.5%	97.5%	2.5%	97.5%
RoV	19	2.662	6.431	2.525	5.907
MS2	13	2.544	8.119	2.516	12.499
HAdV	10	1.027	5.200	0.932	4.059
H. pylori	9	2.023	5.734	2.054	8.467
E. coli	10	0.887	3.734	0.886	3.609
A. castellanii	9	2.711	10.990	2.627	34.465

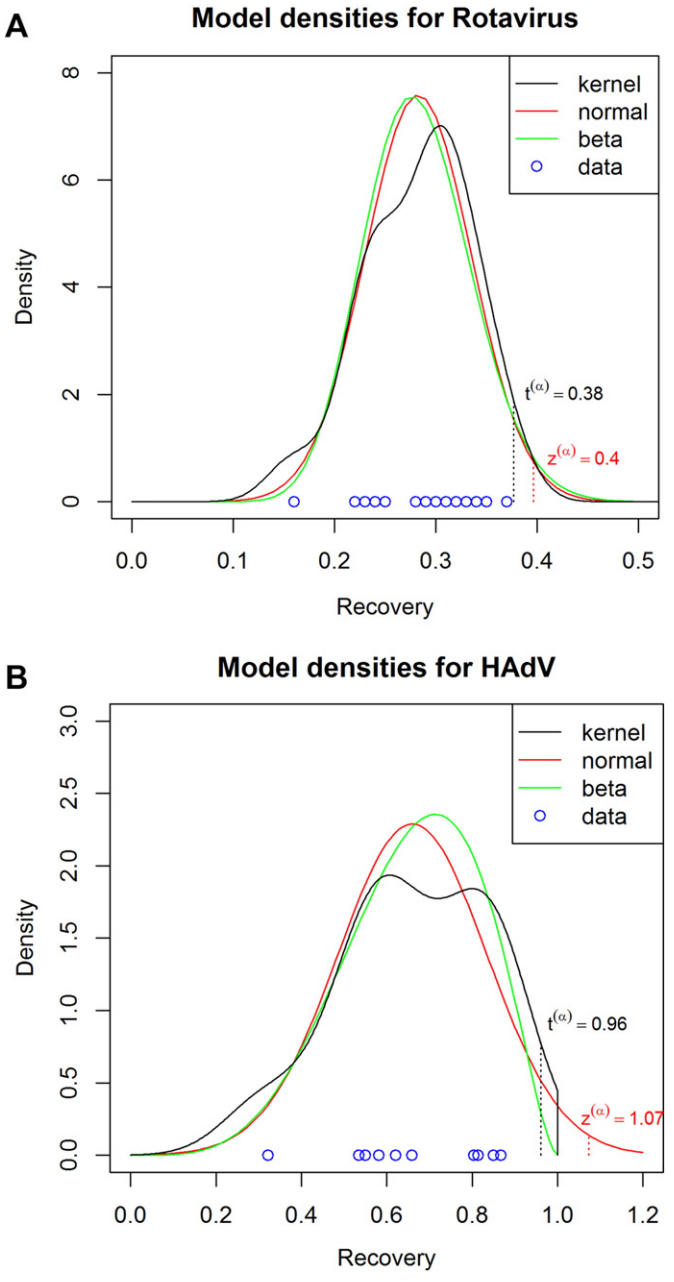


Fig. 1. A: Rotavirus recoveries (dots over the axis) are similarly fitted by the density estimations: normal, beta and kernel. The similar PIs in Table 5 can be explained by the symmetry of the distributions and the similar tails of the 3 models. Notice the value of the quantiles $t^{(\alpha)}$ and $z^{(\alpha)}$ associated to the upper bounds of the bootstrap-t and normal prediction intervals with $1 - \alpha = 95\%$ confidence. B: In contrast, in HAdV, the 3 models show a different fit on the upper right tail of the distribution. The discordance between the intervals for HAdV in Table 5 can be explained by taking into account that $t^{(\alpha)}$ and $z^{(\alpha)}$ are computed precisely on the tails of each distribution.

between bacteria and between viruses support the theory that *E. coli* and HAdV are suitable indicators for bacterial and viral contamination, respectively.

A ratio between the logarithm of the quantification between q(RT)PCR and infectivity suggests that the difference in quantification for these viruses was related to the proportion of non-infectious particles that may be produced in the cell lines where they have been cultured and the different sensitivities of the assays. The number of HAdVs detected in water using qPCR are typically 1 to 2 logs higher than estimates using culture-based methods (Rames et al., 2016). Moreover, HAdV and RoV are recognized to be resistant to pH, and their infectivity is not affected by the acidification of the sample in the SMF protocol (Attoui et al., 2012; Harrach et al., 2012). Although MS2 shows the smallest difference in quantification, the percentage of recoveries differ between q(RT)PCR and plaque-forming units. However, for an enveloped virus such as BVDV, either the recovery or the quantification was higher in q(RT)PCR (ratio of 2.03), which could be due to the acid pH (3.5) treatment (for approximately 16 h) that is used in the SMF protocol. In general, the sensitivity of infectivity assays has been traditionally considered to be lower in comparison with PCR techniques (Amini and Kraatz, 2014).

It is important to note that the model captures the random character of the unknown recoveries, but does not capture the random character of the concentrations in the sampled region. Therefore, the fitted error in the expression above refers to the technical error in the measurements but not the actual distribution of the organisms' concentrations in the water.

We strongly recommend using the extrapolation method with samples previously spiked with a surrogate virus or process control, such as the MS2 used in this study. We recommend verification of the recovery obtained with this surrogate, which might be between 10% and 38% (within 2 standard deviations of the mean of recovery) and is an interval that allows us to describe the variability and uncertainty of the SMF method in our laboratory. Another way to determine the recovery percentage in each laboratory is to estimate the mean and standard deviation of the surrogate under the particular laboratory conditions. Although variability is an intrinsic characteristic of each variable, the uncertainty introduces subjective "variability" into the variable (Vose, 2008). In addition, it may be increased or decreased by the expertise of the operator developing the SMF, the equipment, and the reagents. These may cause differences in the recoveries between operators or laboratories and should be taken into account.

The SMF recoveries are susceptible to improvement, and it is important to minimize the attachment of flocs to the lateral wall of the bucket after the sedimentation step and the loss of small pieces of flocs in the decanting process after the centrifugation step at 8000 rpm for 30 min. Although the use of non-adherent buckets may increase the recovery efficacy, this will make the SMF method more expensive. The fungible materials per sample were estimated to be low cost and can be disinfected or reused for other purposes or recycled.

The parameters that provide variability and uncertainty in HAdV and *E. coli* will be very useful in future studies. Under controlled conditions, the variability between the samples for the percentage of recovery in HAdV (CI 95%: 53.5–78.5%) includes the results of previous studies

with SMF (Calgua et al., 2008; Rusiñol et al., 2014, 2015). Therefore, we suggest modelling the recovery with the purpose of having a better approach for the real risk of the presence of the microorganism in water.

While the risk to a population is dictated by the frequency of contamination and the distribution of the dose, the probability of infection of an individual is ultimately based on the number of pathogens ingested (Ross, 2008). Errors in the precision of the quantification can underestimate the real concentration. Therefore, an extrapolation method that permits estimation of the real concentration of microorganisms in water samples is important in obtaining a better approach for future QMRA.

The distribution of the recovery under the controlled condition does not suggest that the distribution of the microorganisms in the environment occurs in the same way. It is important to consider that sampling methods of water in the environment have always been a limitation and require more work to determine the right way to describe the distribution of microorganisms in the environment (Petterson et al., 2015). Microbial water quality often varies rapidly and over a wide range. Short-term peaks in pathogen concentrations may increase disease risk considerably and may trigger outbreaks of water-borne disease, and furthermore, by the time the microbial contamination is detected, many people may have been exposed (Gorchev and Ozolins, 2011).

In summary, the low cost, repeatability, low variability, and applicability to the methods described for the simultaneous concentration of a diversity of microorganisms support SMF as a useful tool for the control of water safety. In addition, the possibility of obtaining intervals, which allows the prediction of the actual amount of microorganisms in the samples, including the uncertainty of the method, shows that SMF is an efficacious and efficient method for concentration and should be considered a robust procedure for evaluating the microbiological quality of water and the associated public health risk.

5. Conclusions

SMF can be used to efficiently and simultaneously concentrate viruses, bacteria and protozoa with repeatable results.

Statistically significant positive correlations were found between the recoveries of the bacteria evaluated, *E. coli* and *H. pylori*. Although the correlation between the recoveries of the viruses was not statistically significant, a positive correlation between them shows that HAdV is a suitable indicator for viral contamination.

Quantification by q(RT)PCR and infectivity methods shows ratios that suggest similar recoveries for HAdV-35 and RoV. They may be used indistinctly to evaluate these microorganisms with an SMF method of concentration.

The quantification of BVDB, which is sensitive to pH in the process of flocculation, is more efficiently conducted using q(RT)PCR than infectivity assays.

The estimation of the inoculums using q(RT)PCR quantification and the 95% bootstrap PI using the sample of the recovery estimates for each microorganism permits the acquisition of intervals that predict the real concentration of pathogens or indicators and may be used as a measure of uncertainty in QMRA studies.

Funding information

This study was funded by the Water Challenges for a Changing World Joint Programming Initiative (W2013-095-C03-01), the Spanish Ministry of Economy and Competitiveness (MINECO — AGL2014-55081-R) and the Grup de Recerca Consolidat: Virus, bacteris i protozous d'interès en salut pública, aigua y aliments (Generalitat de Catalunya, Virbap — 2014-SGR-914).

Acknowledgements

Eloy Gonzales-Gustavson would like to acknowledge the Presidente de la República scholarship from the Peruvian Government, which supported a PhD grant.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.mimet.2017.01.006.

References

Amini, K., Kraatz, H.-B., 2014. Recent advances and developments in monitoring biological agents in water samples. Rev. Environ. Sci. Biotechnol. 14:23–48. http://dx.doi.org/10.1007/s11157-014-9351-5.

Anonymous, 1995. ISO 10705-1: Water Quality—Detection and Enumeration of Bacteriophages-Part 1: Enumeration of F-specific RNA Bacteriophages. International Organisation for Standardisation, Geneva, Switzerland.

Attoui, H., Mertens, P., Becnel, J., Belaganahali, S., Betgoin, M., Brussaard, C., Chappell, J., Ciarlet, M., de Vas, M., Dermody, T., Dormitzer, P., Duncan, R., Fang, Q., Graham, R., Guglielmi, K., Harding, R., Hilman, B., Makkay, S., Marzachì, C., Matthijnssens, J., Milne, R., Mohd Jaafar, F., Mori, H., Noordeloos, A., Omura, T., Patton, J., Rao, S., Mann, M., Stolts, D., Suzuki, N., Upadhyaya, N., Wei, C., Zhou, H., 2012. Family reoviridae. In: King, A., Adams, M., Carstens, E., Lefkowitz, E. (Eds.), Virus Taxonomy, Classification and Nomenclature of Viruses. Elsevier Academic Press, Davis, pp. 541–637.

Bai, C., Bickel, P., Olshen, R., Marcus, M.B., 1990. Hyperaccuracy of bootstrap based prediction. In: Eberlein, E., Kuelbs, J. (Eds.), Probability in Banach Spaces VII Proceedings of the Seventh International Conference. Birkhauser, Boston, Cambridge, Massachusetts, pp. 31–42.

Bofill-Mas, S., Hundesa, A., Calgua, B., Rusinol, M., Maluquer de Motes, C., Girones, R., 2011. Cost-effective method for microbial source tracking using specific human and animal viruses. J. Vis. Exp. http://dx.doi.org/10.3791/2820.

Bofill-Mas, S., Rusiñol, M., Fernandez-cassi, X., Carratalà, A., Hundesa, A., Girones, R., 2013. Quantification of human and animal viruses to environmental samples. Biomed. Res. Int.:2013 http://dx.doi.org/10.1155/2014/675857.

Calgua, B., Mengewein, A., Grunert, A., Bofill-Mas, S., Clemente-Casares, P., Hundesa, A., Wyn-Jones, A.P., Lopez-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. http://dx.doi.org/10.1016/j.jviromet.2008.08.003.

Calgua, B., Barardi, C.R.M., Bofill-Mas, S., Rodriguez-Manzano, J., Girones, R., 2011. Detection and quantitation of infectious human adenoviruses and JC polyomaviruses in water by immunofluorescence assay. J. Virol. Methods 171:1–7. http://dx.doi.org/10.1016/j.jviromet.2010.09.013.

Calgua, B., Fumian, T., Rusinol, 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. http://dx.doi.org/10.1016/j.watres.2013.02.043.

Calgua, B., Rodriguez-Manzano, J., Hundesa, A., Sunen, E., Calvo, M., Bofill-Mas, S., Girones, R., 2013b. New methods for the concentration of viruses from urban sewage using quantitative PCR. J. Virol. Methods 187:215–221. http://dx.doi.org/10.1016/j.jviromet.2012.10.012.

Calgua, B., Carratala, A., Guerrero-Latorre, L., de Abreu Correa, A., Kohn, T., Sommer, R., Girones, R., 2014. UVC inactivation of dsDNA and ssRNA viruses in water: UV fluences and a qPCR-based approach to evaluate decay on viral infectivity. Food Environ. Virol.:260–268 http://dx.doi.org/10.1007/s12560-014-9157-1.

Estes, M., Greenberg, H., 2013. Rotaviruses. In: Knipe, D., Howley, P. (Eds.), Fields Virology. Lippincott Williams & Wilkins, Philadelphia, pp. 1347–1401.

Eusebi, L.H., Zagari, R.M., Bazzoli, F., 2014. Epidemiology of *Helicobacter pylori* infection. Helicobacter 19:1–5. http://dx.doi.org/10.1111/hel.12165.

Gascón, J., 2006. Epidemiology, etiology and pathophysiology of traveler's diarrhea. Digestion 73:102–108. http://dx.doi.org/10.1159/000089785.

Gorchev, H.G., Ozolins, G., 2011. WHO guidelines for drinking-water quality. WHO Chron. 38:104–108. http://dx.doi.org/10.1016/S1462-0758(00)00006-6.

Harrach, B., Benkö, M., Both, G., Brown, M., Davidson, A., Echavarría, M., Hess, M., Jones, M., Kajon, A., Lehmkuhl, H., Mautner, V., Mittal, S., Wadell, G., 2012. Family Adenoviridae. In: King, A., Adams, M., Carstens, E., Lefkowitz, E. (Eds.), Virus Taxonomy, Clasification and Nomenclature of Viruses. Elsevier Academic Press, San Diego, pp. 125–141.

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

Hill, V.R., Polaczyk, A.L., Hahn, D., Narayanan, J., Cromeans, T.L., Roberts, J.M., Amburgey, J.E., 2005. Development of a rapid method for simultaneous recovery of diverse microbes in drinking water by ultrafiltration with sodium polyphosphate and surfactants. Appl. Environ. Microbiol. 71:6878–6884. http://dx.doi.org/10.1128/AEM.71.11.6878-6884.2005.

Hill, V.R., Narayanan, J., Gallen, R.R., Ferdinand, K.L., Cromeans, T., Vinjé, J., 2015. Development of a nucleic acid extraction procedure for simultaneous recovery of DNA and

RNA from diverse microbes in water. Pathogens 4:335–354. http://dx.doi.org/10.3390/pathogens4020335.
Johnson, C.H., Rice, E.W., Reasoner, D.J., 1997. Inactivation of *Helicobacter pylori* by chlorination. Appl. Environ. Microbiol. 63, 4969–4970.
Khan, I.U.H., Gannon, V., Kent, R., Koning, W., Lapen, D.R., Miller, J., Neumann, N., Phillips, R., Robertson, W., Topp, E., van Bochove, E., Edge, T.A., 2007. Development of a rapid quantitative PCR assay for direct detection and quantification of culturable and non-culturable *Escherichia coli* from agriculture watersheds. J. Microbiol. Methods 69: 480–488. http://dx.doi.org/10.1016/j.mimet.2007.02.016.
Lambrecht, E., Baré, J., Chavatte, N., Bert, W., Sabbe, K., Houf, K., 2015. Protozoan cysts act as survival niche and protective shelter for foodborne pathogenic bacteria. Appl. Environ. Microbiol. 81. http://dx.doi.org/10.1128/AEM.01031-15 (AEM.01031-15).
Losurdo, M., Mari, V., Lucente, M.S., Colaianni, M.L., Padalino, I., Cavaliere, N., Buonavoglia, C., Decaro, N., 2015. Development of a TaqMan assay for sensitive detection of all pestiviruses infecting cattle, including the emerging HoBi-like strains. J. Virol. Methods 224:77–82. http://dx.doi.org/10.1016/j.jviromet.2015.08.013.
MacLachlan, J., Dubovi, E., 2011. Fenner's Veterinary Virology. fourth ed. Academic Press, San Diego http://dx.doi.org/10.1016/B978-0-12-375158-4.00030-4.
Madigan, M., Martinko, J., Bender, K., Buckley, D., Stahl, D., 2014. Brock Biology of Microorganisms. 14th ed. Pearson, Harlow.
Marciano-Cabral, F., Cabral, G., 2003. *Acanthamoeba* spp. as agents of disease in humans. Clin. Microbiol. Rev. 16:273–307. http://dx.doi.org/10.1128/CMR.16.2.273-307.2003.
Melgaço, F.G., Victoria, M., Correa, A.A., Ganime, A.C., Malta, F.C., Brandao, M.L.L., de Mello Medeiros, V., de Oliveira Rosas, C., Bricio, S.M.L., Miagostovich, M.P., 2016. Virus recovering from strawberries: evaluation of a skimmed milk organic flocculation method for assessment of microbiological contamination. Int. J. Food Microbiol. 217: 14–19. http://dx.doi.org/10.1016/j.ijfoodmicro.2015.10.005.
Mojirsheibani, M., 1998. Iterated bootstrap prediction intervals. Stat. Sin. 8, 489–504.
Moreno, Y., Ferrús, M.A., 2012. Specific detection of cultivable *Helicobacter pylori* cells from wastewater treatment plants. Helicobacter 17:327–332. http://dx.doi.org/10.1111/j.1523-5378.2012.00961.x.
OIE, 2015. Bovine viral diarrhea. Manual of Diagnostic Tests and Vaccines for Terrestrial Animals. Office International des Epizooties, Paris, France, pp. 1–22.
Otto, P.H., Reetz, J., Eichhorn, W., Herbst, W., Elschner, M.C., 2015. Isolation and propagation of the animal rotaviruses in MA-104 cells—30 years of practical experience. J. Virol. Methods 223:88–95. http://dx.doi.org/10.1016/j.jviromet.2015.07.016.
Petterson, S., Grøndahl-Rosado, R., Nilsen, V., Myrmel, M., Robertson, L.J., 2015. Variability in the recovery of a virus concentration procedure in water: implications for QMRA. Water Res. 87:79–86. http://dx.doi.org/10.1016/j.watres.2015.09.006.
Pina, S., Puig, M., Lucena, F., Jofre, J., Girones, R., 1998. Viral pollution in the environment and in shellfish: human adenovirus detection by PCR as an index of human viruses. Appl. Environ. Microbiol. 64, 3376–3382.
Qvarnstrom, Y., Visvesvara, G.S., Sriram, R., Da Silva, A.J., 2006. Multiplex real-time PCR assay for simultaneous detection of *Acanthamoeba* spp., *Balamuthia mandrillaris*, and *Naegleria fowleri*. J. Clin. Microbiol. 44:3589–3595. http://dx.doi.org/10.1128/JCM.00875-06.
R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (URL https//www.R-project.org/).
Rames, E., Roiko, A., Stratton, H., Macdonald, J., 2016. Technical aspects of using human adenovirus as a viral water quality indicator. Water Res. 96:308–326. http://dx.doi.org/10.1016/j.watres.2016.03.042.
Ross, T., 2008. Microbial ecology in food safety risk assessment. In: Schaffner, D. (Ed.), Microbiological Risk Assessment of Foods. ASM Press, Washington DC, pp. 51–98.
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. http://dx.doi.org/10.1016/j.watres.2014.04.013.
Rusiñol, M., Fernandez-Cassi, X., Timoneda, N., Carratala, 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. Manag. 159:58–67. http://dx.doi.org/10.1016/j.jenvman.2015.05.019.
Santiago, P., Moreno, Y., Ferrús, M.A., 2015. Identification of viable *Helicobacter pylori* in drinking water supplies by cultural and molecular techniques. Helicobacter 20: 252–259. http://dx.doi.org/10.1111/hel.12205.
USEPA, 2005. Method 1623: Cryptosporidium and Giardia in Water by Filtration/IMS/FA. EPA 815-R-05-002. Off. Water, Washington, DC.
van Duin, J., Olsthoorn, R.C.L., 2012. Family Leviviridae. In: King, A., Adams, M., Carstens, E., Lefkowitz, E. (Eds.), Virus Taxonomy, Classification and Nomenclature of Viruses. Elsevier Academic Press, Davis, pp. 1035–1043.
Victoria, M., Fumian, T.M., Rocha, M.S., Dalmao, F., Leite, J.P.G., Girones, R., Miagostovich, M.P., 2014. Gastroenteric virus dissemination and influence of rainfall events in urban beaches in Brazil. J. Appl. Microbiol. 117:1210–1218. http://dx.doi.org/10.1111/jam.12592.
Vose, D., 2008. Risk Analysis a Quantitative Guide. third ed. John Wiley and Sons, Wiltshire, England.
Wold, W.S., Horwitz, M., 2013. Adenoviruses. In: Knipe, D.M., Howley, P.M. (Eds.), Fields Virology. Lippincott Williams & Wilkins, Philadelphia, pp. 1732–1767.
World Health Organization, 2014. Preventing Diarrhoea Through Better Water, Sanitation and Hygiene. pp. 1–32 (Online).
Zeng, S.Q., Halkosalo, A., Salminen, M., Szakal, E.D., Puustinen, L., Vesikari, T., 2008. One-step quantitative RT-PCR for the detection of rotavirus in acute gastroenteritis. J. Virol. Methods 153:238–240. http://dx.doi.org/10.1016/j.jviromet.2008.08.004.

Journal of Microbiological Methods 142 (2017) 96

Contents lists available at ScienceDirect

Journal of Microbiological Methods

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

Corrigendum

Corrigendum to "Characterization of the efficiency and uncertainty of skimmed milk flocculation for the simultaneous concentration and quantification of water-borne viruses, bacteria and protozoa" [J. Microbiol. Methods 134 (2017) 46–53]

Eloy Gonzales-Gustavson[a], Yexenia Cárdenas-Youngs[a], Miquel Calvo[a], Marcelle Figueira Marques da Silva[a], Ayalkibet Hundesa[a], Inmaculada Amorós[b], Yolanda Moreno[b], Laura Moreno-Mesonero[b], Rosa Rosell[c], Llilianne Ganges[c], Rosa Araujo[a], Rosina Girones[a,*]

[a] Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, Av. Diagonal 643, 08028 Barcelona, Catalonia, Spain
[b] Instituto de Ingeniería del Agua yMedio Ambiente, Ciudad Politécnica de la Innovación, Ed. 8G, Acceso D, planta 2, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia, Spain
[c] IRTA, Centre de Recerca en Sanitat Animal (CReSA, IRTA-UAB), Campus de la Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

The authors regret

Table 2: HAdV IFA (second row) min and max says 8.1 and 49.8 and must say 43.5 and 83.7

In Materials and methods, Section 2.1 Microorganism stock and cell lines: line 8 says NDBK and must say MDBK.

The authors would like to apologize for any inconvenience caused.

DOI of original article: http://dx.doi.org/10.1016/j.mimet.2017.01.006
* Corresponding author.
E-mail address: rgirones@ub.edu (R. Girones).

http://dx.doi.org/10.1016/j.mimet.2017.07.016

Available online 09 August 2017

4.2 Artículo 2: Caracterización de la concentración de norovirus y adenovirus en agua regenerada y evaluación del riesgo de su uso para irrigar lechugas en Cataluña

Quantitative risk assessment for the use of reclaimed water to irrigate lettuce in Catalonia

Gonzales-Gustavson, E.; Rusiñol M; Medema G; Calvo, M; R. Girones.

El agua residual tratada puede ser una importante fuente de recursos en regiones con escasas precipitaciones en el mundo. Su utilización en agricultura depende de los riesgos de salud que son medidos a través de la presencia de bacterias indicadoras de contaminación fecal. Sin embargo, está demostrado que las ausencias de estos indicadores comunes de contaminación no están relacionadas con la presencia de virus, los principales agentes causantes de enfermedades transmitidas por el agua. En este estudio, se caracterizó durante un año la concentración de HAdV y NoV GII, los virus más frecuentes en agua residual, en dos plantas de tratamiento de agua residual (WWTPs) situadas en Cataluña que disponen de tratamientos terciarios diferentes (lagunaje natural vs tratamiento tecnificado convencional). El objetivo principal fue realizar un estudio de QMRA para estimar el riesgo para la salud asociado con la ingestión de lechugas irrigadas con efluente terciario proveniente de las dos WWTPs. El primero es un escenario conservador desarrollado con un modelo de dosis-respuesta que incluye el efecto de agregación viral para NoV GII y la probabilidad de enfermedad condicionada a infección que es dependiente de la dosis. El segundo es un escenario desfavorable sin agregación para NoV GII y con valores fijos de probabilidad de enfermedad condicionada a infección. Los resultados obtenidos muestran que la carga de enfermedad ocasionada por NoV GII y HAdV por el consumo de lechugas irrigadas con efluente terciario no llega a alcanzar los límites de seguridad establecidos por la WHO de 10^{-6} DALYs a excepción del escenario conservador con NoV GII en la WWTP con lagunaje natural como tratamiento terciario. Esta planta mostró en promedio una mejor reducción, pero con una mayor variabilidad en el tratamiento con respecto a la de tratamiento tecnificado convencional, sin embargo, el lagunaje es capaz de tratar únicamente el 10% del volumen total de agua residual que reciben en la planta. El análisis de sensibilidad demostró que los parámetros de entrada que más

afectan la variabilidad del modelo son la concentración viral y la reducción en las WWTPs. Para lograr satisfacer las recomendaciones de WHO son necesarias reducciones en promedio de entre 4 a 9 logaritmos en el escenario conservador y de entre 7 y 15 en el escenario desfavorable o estimar concentraciones puntuales de 0.5 CG/ml en agua regenerada en el escenario desfavorable. Los resultados de este análisis sugieren que la utilización del agua regenerada analizada en este estudio podría acarrear riesgo de enfermedad viral y es necesario considerar la adición de barreras o tratamientos de desinfección para reducir su carga en las WWTPs y así asegurar la calidad de los vegetales irrigados con esta agua.

Quantitative risk assessment of norovirus and adenovirus for the use of reclaimed water to irrigate lettuce in Catalonia

Gonzales-Gustavson, E.; Rusiñol, M; Medema, G; Calvo, M; Girones, R.

Abstract

Wastewater is an important resource in water-scarce regions of the world, and its use in agriculture requires the guarantee of acceptable risk levels in public health. The presence of fecal bacteria, indicators of contamination, does not correlate with the presence of viruses, which are the main potential health risks transmitted through water. Using viral pathogens as indicators could complement the use of fecal indicator bacteria in the evaluation of water quality. In this study, we characterized the concentration of human adenovirus (HAdV) and norovirus genogroup II (NoV GII), the most important human viruses found in wastewater, in two wastewater treatment plants (WWTPs) that use different tertiary treatments (natural wetland vs conventional UV, Cl and Actiflo treatments) for a year in Catalonia. The main objective of this study was to develop a quantitative microbial risk assessment to estimate the health risk associated with the ingestion of lettuce irrigated with tertiary effluents from these WWTPs. The results show that the disease burden of NoV GII and HAdV for the consumption of lettuce irrigated with tertiary effluent from either WWTP was higher than the WHO recommendation of 10^{-6} DALYs for both viruses. The WWTP with natural wetland showed a higher viral reduction on average (12.2 and 8.9 logs for NoV GII and HAdV, respectively) than conventional treatment (5.2 and 3.7 logs) but a higher variability than the conventional WWTP. Sensitivity analysis demonstrated that the input parameters used to estimate the viral reduction by treatment and viral concentrations accounted for much of the model output variability. The estimated reductions required to reach the WHO recommended levels in tertiary effluent depended mainly on the treatments developed in the WWTPs, and additional average

reductions are necessary (in WWTP 1: 3 logs for HAdV and 4.8 logs for NoV GII; and in WWTP 2: 2.6 logs for HAdV and 2.4 for NoV GII). This recommendation would be achieved with an average quantification of 0.5 genome copies per ml in reclaimed water. The results suggest that the analyzed reclaimed water would require an extra disinfection treatment to achieve acceptable risk in the irrigation of vegetables with reclaimed water.

Introduction

Reuse of wastewater for agricultural irrigation is being widely implemented because water scarcity is reported in nearly all river basins in the Mediterranean area. Wastewater is often a reliable year-round source of water, and it contains necessary nutrients for plant growth. Wastewater needs to be treated to produce reclaimed water for it to be used for irrigation (EU, 2016; Sanz and Gawlik, 2014). For example, Spain uses 71% of its total volume of reclaimed water for agricultural irrigation (Iglesias et al., 2010). Reclaimed water is also used for urban, industrial, recreational and environmental activities.

The use of reclaimed water in Spain is regulated under the Real Decreto 1620/2007. This regulation establishes the minimum acceptable safety limits for each type of use in Spain, including agricultural irrigation. These limits include the levels of intestinal nematode eggs, *Escherichia coli*, suspended solids and turbidity (Boletin Oficial del Estado, 2007), but this regulation does not include addressing the acceptable levels of viruses. Food crops irrigated with untreated or poorly treated water are a main source of viruses in foodborne outbreaks. Bacterial or parasite indicators are poorly related to the presence of human enteric viruses (Petterson et al., 2001).

The control of the microbiological quality of reclaimed water in wastewater treatment plants (WWTPs) is currently based on the levels of fecal indicator bacteria (FIB), which include fecal coliforms, *Escherichia coli* and enterococci. However, the FIB displays behave differently than

enteric viruses in wastewater and aquatic environments, where these bacteria are more susceptible to water treatments and environmental conditions (McMinn et al., 2017). Despite viruses not being the only type of waterborne pathogen that may cause disease, the risk of illness from viruses is 10 – 10000 times greater than that from bacteria at a similar level of exposure (Haas et al., 1993). For that reason, the unique evaluation of the FIB underestimates the public health risk of enteric viruses.

The most effective means of consistently ensuring safety in the agricultural application of wastewater is through the use of a comprehensive risk assessment and risk management approach that encompasses all steps in the process from waste generation to the treatment and use of wastewater to product use or consumption (WHO, 2006). Quantitative Microbial Risk Assessments (QMRAs) translate the pathogen dose that the consumer is exposed to in a particular scenario into probabilities of infection and illness that can be compared against a tolerable disease burden. Disability-Adjusted Life Years (DALYs) are the recommended metric in the WHO guidelines for the overall community health burden, and the tolerable value is 10^{-6} DALY loss per person per year (pppy) (WHO, 2006).

Among the viruses of fecal origin that are present in reclaimed water, norovirus (NoV) is the main cause of viral gastroenteritis in people of all ages worldwide and is replacing rotavirus as the predominant gastrointestinal pathogen in children. This virus is often found in wastewater and selected as reference virus in QMRAs in a broad variety of scenarios, including exposure to irrigated crops (Allende and Monaghan, 2015; Barker, 2014; Mara and Sleigh, 2010; Mok et al., 2014; Owusu-Ansah et al., 2017; Sales-Ortells et al., 2015). Previous epidemiological studies have demonstrated that NoV genogroup II (NoV GII), including the genotypes GII.2, GII.3, GII.4, and GII.6, is the main cause of endemic persistence and recent large outbreaks of gastroenteritis. Furthermore, another genotype, the GII.P17-GII.17 virus, emerged in 2013 and is spreading as fast as GII.4 (Kobayashi et al., 2016).

Another virus transmitted by contaminated food and water is human adenovirus (HAdV), which is highly prevalent and resistant to sewage treatment (Adefisoye et al., 2016; Calgua et al., 2013b; Grøndahl-Rosado et al., 2014). This virus has been recommended as an indicator for human contamination in water (Albinana-Gimenez et al., 2009; Pina et al., 1998; Rusiñol et al., 2015; Wyn-Jones et al., 2011). However, little scientific information is available about the transmission of HAdV through vegetables. HAdVs can cause an array of clinical diseases, including conjunctivitis, gastroenteritis, myocarditis, and pneumonia (Ghebremedhin, 2014). However, HAdVs and NoV rarely cause serious illness or death although infants and people with weakened immune systems or existing respiratory or cardiac disease are at higher risk of developing severe disease.

This study characterizes the HAdV and NoV GII viral concentrations in reclaimed water based on q(RT)PCRs using a mathematical approach that models the variability of the viral load before and after treatment and its reduction in WWTPs. Moreover, we assess the health risk associated with the consumption of lettuce irrigated with reclaimed water from two WWTPs with different tertiary treatments: conventional and wetland.

Methods

Study site description

Two WWTPs located in the northeast of Spain were selected. WWTP 1 was designed to treat wastewater from two million inhabitants with a flow capacity of 420,000 m^3/day. WWTP 2 was designed to treat wastewater from 112,000 inhabitants with a flow capacity of 30,000 m^3/day. Both WWTPs have conventional secondary treatments that consist of sedimentation and activated sludge. WWTP 1 has a tertiary treatment, with a design capacity of 3.25 m^3/s, that consists of chlorination, flocculation (Actiflo®) and UV treatment. WWTP 2 introduces 10% of the secondary treatment water into a wetland that is located next to the WWTP as tertiary treatment. In both

WWTPs, part of the reclaimed water is generated by local people to irrigate the vegetables of small farms.

Sampling, concentration and molecular quantification:

For both WWTPs, monthly samples were taken of raw sewage after secondary treatment and after tertiary treatment for one year. At each site, 500 ml and 10 L of raw and treated wastewater, respectively, were collected. Viruses in these samples were concentrated using the skimmed milk flocculation (SMF) method for raw (Calgua et al., 2013a) and treated water (Calgua et al., 2008). Viral nucleic acids were extracted using a QIAmp Viral RNA kit (Qiagen, Inc., Valencia, CA) following the manufacturer's instructions. Samples were tested for the viral pathogens HAdV (Hernroth et al., 2002) and NoV GII (Kageyama et al., 2003) using real-time qPCR and RT-qPCR, respectively. Undiluted and 10-fold diluted samples of the nucleic acid extracts were analyzed in duplicate, including the concentrates from negative control buckets. All of the q(RT)PCR assays included four non-template controls to demonstrate that the reaction mix itself did not produce fluorescence. The virus standards were prepared using synthetic gBlocks® Gene Fragments (IDT®) and quantified with a Qubit® fluorometer (Thermo Fisher Scientific). Ten-fold dilutions were used to prepare samples with concentrations ranging from 10^0 to 10^7 copies per reaction. The MS2 virus was spiked into and monitored in all the samples as a control to ensure the efficacy of the laboratory procedure.

Quantitative microbial risk assessment

The QMRA was constructed for lettuce consumption patterns to determine the DALYs following the steps suggested by the WHO guidelines (WHO, 2016) as described in the following paragraphs.

Hazard identification:

The reference pathogens HAdV and NoV GII were selected to provide a model to describe the viral risk of waterborne transmission through contaminated vegetables. HAdV is a double-stranded DNA

virus that belongs to the *Adenoviridae* family. NoV is a single-stranded RNA that belongs to the *Caliciviridae* family. Both viruses were chosen because they are the most common cause of gastroenteritis illness in Catalonia; additionally, they are commonly found in water, are resistant to environmental degradation and differ in their sensitivity to water treatment processes such as UV light exposure (Hijnen et al., 2006; Rusiñol et al., 2015, 2014).

Exposure assessment:

To estimate the concentrations of these reference viruses, we fitted the monitoring data to a negative binomial distribution. The distributions describe the variability in virus concentration in the water source. Virus reduction by a water treatment (π) was estimated stochastically in each WWTP using a Beta distribution based on a method described previously (Teunis et al., 1999, 2009). This method allows an unequal number of samples before and after treatment to be used with the advantage of including zero counts in the model.

The negative binomial distribution for raw sewage is described as

$$g(n, V|r, \lambda) = \frac{\Gamma(n+r)}{n! \times \Gamma(r)} \times \frac{(\lambda \times V)^n}{(1+\lambda \times V)^{n+r}} \quad (1)$$

where n is the number of viruses in raw sewage, V is the volume of water evaluated and λ and r are the scale and shape parameters of the gamma distribution, respectively.

The distribution of the number of viruses in a sample after treatment C_{eff} is

$$h(k, W|\lambda, \rho, \alpha, \beta) = (\lambda \times W)^k \frac{\Gamma(r+k)}{k!\Gamma(r)} \times \frac{\Gamma(\alpha+\beta) \times \Gamma(\alpha+k)}{\Gamma(\alpha) \times \Gamma(\alpha+\beta+k)} \times_2 F_1(k+r, \alpha+k, \alpha+\beta+k, -\lambda \times W)$$

(2)

where *k* is the number of viruses after treatment, W is the volume of water evaluated, α and β are the shape parameters of the Beta distribution (π), which expresses the reduction in the number of viruses due to the treatment, and ${}_2F_1$ is the Gaussian hypergeometric function. The parameters

were estimated by maximum likelihood following the method described by Teunis et al. (Teunis et al., 1999 and 2009) for unpaired samples.

Based on the suggestion of previous studies, the viral enumeration data were also corrected in the assessment to account for viral loss during the concentration procedure (Petterson et al., 2015). The concentration was corrected with a Beta distribution, with the recoveries previously described specifically for the SMF method used to concentrate the viruses (Table 1) (Gonzales-Gustavson et al., 2017).

The scenario modeled in this study involved the consumption of lettuce irrigated with tertiary-treated water. This vegetable was chosen because lettuce potentially protects viruses from light and desiccation, thus enhancing pathogenic persistence (Petterson et al., 2001). Moreover, leafy greens, such as lettuce, are prone to contamination with pathogens as they have large surface areas, are grown in close proximity to soil, are irrigated intensively and are mainly consumed raw (De Keuckelaere et al., 2015). This paper considered only overhead sprinkler irrigation because it is the method used in the field. The transfer of viruses to lettuce by irrigation was described in a previous study (Mok and Hamilton, 2014), and its stochastic description was used here.

The in-field virus decay (Rs) and the inactivation that occurs during storage and transport (Rt) were included in the analysis based on a previous study with HAdV and MS2 (Carratalà et al., 2013) and assumed to be between 1 and 2 $\log_{10}$ in the period between the last irrigation and harvesting and between 0 and 1 $\log_{10}$ during dark storage and transport. Additionally, lettuce washing reduces virus concentrations between 0.1 and 2 $\log_{10}$ and was described here with a PERT distribution (Mok et al., 2014). To estimate the level of exposure, we assumed the daily rate of lettuce consumption in Spain to be lognormal distributed based on the national census of Spain, which described the per capita Spanish consumption of lettuce (Aecosan, 2015). Finally, the daily dose of viruses on lettuce surfaces

(d_s) ingested by consumers in the area where the lettuce irrigated with reclaimed water had been sold was calculated by:

$$d_{NoV} = C_{eff} \times 10^{(-R_s - R_t - R_{wash})} \times V_{surf} \times \frac{1}{\pi_{rec}} \times I \quad (3)$$

where C_{eff} is the concentration in tertiary effluent per ml, V_{surf} represents the clinging of viruses to the lettuce, R_s is the reduction in the number of viruses on the surface due to UV light and high temperatures in the field, R_t is the reduction in the number of viruses between harvest and consumption, R_{wash} is the reduction in the number of surface viruses due to washing with water, π_{rec} is the recovery factor of the concentration method (SMF) and I is the amount of lettuce ingested. The general fitting parameters for the probability distributions are shown in Table 1.

The dose-response models for HAdV were developed based on infectious particles, while the data in this study are qPCR-based. An additional parameter was therefore included to estimate the dose of infectious HAdV (eq. 4): the ratio of infectious particles to genome copies (GC) detected by qPCR (R_{inf}) was based on information published previously (Gonzales-Gustavson et al., 2017; Rames et al., 2016). For NoV, both dose-response data and wastewater data are q(RT)PCR-based, so no correction was needed.

$$d_{HAdV} = C_{eff} \times 10^{(-R_s - R_t - R_{wash} - R_{inf})} \times V_{surf} \times \frac{1}{\pi_{rec}} \times I \quad (4)$$

Table 1: Exposure assessment inputs, units, distributions and parameter values, and references

Model inputs	Notation	Units	Distribution	Source
Recovery HAdV	π_{rec}	proportion	Beta (52.62, 27.07)	(Gonzales-Gustavson et al., 2017)
Recovery NoV GII	π_{rec}	proportion	Beta (161, 235)	(Gonzales-Gustavson et al., 2017)
Water that clings to lettuce surface through sprinkler irrigation	V_{surf}	ml/g	Lognormal3 (-4.57, 0.5, 0.006)	(Mok and Hamilton, 2014)
In-field reduction of surface virus	R_s	$\log_{10}$ units	Uniform (1, 2)	(Carratalà et al., 2013)
Reduction in viruses during transport and storage	R_t	$\log_{10}$ units	Uniform (0, 1)	(Carratalà et al., 2013)
Reduction in surface viruses due to washing	R_{wash}	$\log_{10}$ units	PERT (0.1, 1, 2)	(Mok et al., 2014)
Daily consumption of lettuce	Ing	g pppd	Lognormal (20.72, 26.35) (inf=0, sup=120)	(Aecosan, 2015)
Ratio of infectious particles (only for HAdV)	R_{inf}	$\log_{10}$ units	Uniform (1, 2)	(Rames et al., 2016)
Disease burden per case for HAdV	*DBPC*	DALY	Uniform (0.0481, 0.0587)	(Canada, 2010)
Susceptibility fraction for HAdV	f_s	proportion	Uniform (0.8, 1)	(Mok et al., 2014)

Distribution parameters are Beta (shape parameter α, shape parameter β); Lognormal3 (meanlog, sdlog, threshold); PERT (min, mode, max); Uniform (min, max).

Health effects/dose-response assessment

Dose-response models describe the relationship between exposure and the probability of infection and illness. For NoV, the models described by Teunis et al., 2008 were used. They described two models, one for aggregated NoV and one for non-aggregated NoV. We used the dose-response model without aggregation, assuming that WWTPs efficiently eliminated aggregates (eq. 5):

$$P_{inf}(d_s|\alpha,\beta) = 1 - {}_1F_1(\alpha, \alpha+\beta, -d_s) \quad (5)$$

where ${}_1F_1$ is the Kummer confluent hypergeometric function, α and β are the maximum likelihood estimates for non-aggregated NoV with values of 0.04 and 0.05, respectively, and d_s is the dose (Teunis et al., 2008).

The dose-response model described by Teunis et al. (Teunis et al., 2016) was used for HAdV. Only oral inoculation was considered; equation 5 was used, and maximum likelihood estimates for HAdV by the oral inoculation route were 5.11 and 2.8 for α and β, respectively.

The probability of illness given infection ($P_{ill|inf}$) considered in this study was a fixed value described in the literature: 0.5 (Kundu et al., 2013) and 0.7 (Atmar et al., 2014) for HAdV and NoV, respectively. The daily probability of illness (P_{ill}) was calculated by multiplying the probability of infection (P_{inf}) by the conditional probability of illness given infection.

To estimate the annual risk, we consider multiple exposure events to occur randomly in the period when farmers irrigate crops with the effluent during dry months (214 days per year) (Sales-Ortells et al., 2015). The annual probability of illness was estimated using equation

$$P_{ill\ annual} = 1 - \prod_1^{214}\left(1 - \mathrm{Random}(P_{ill})\right) \quad (6)$$

where $\mathrm{Random}(P_{ill})$ is a random sample from the distribution of P_{ill} (Karavarsamis and Hamilton, 2010).

Risk characterization:

Risk characterization was carried out by combining all the information of the hazard identification, exposure assessment and dose-response assessment. We translated the probability of illness into DALYs (pppy) as an annual disease burden output. We estimated the DALYs as:

$$DALY = P_{ill\ annual} \times DBPC \times f_s \quad (7)$$

where $P_{ill\ annual}$ is the annual probability of illness per virus, DBPC is the disease burden (DALYs per case) and f_s is the proportion of the population susceptible to the disease. Since there is no disease burden estimation for either HAdV or NoV in Catalonia, we evaluated two values used previously: a) a mix of Spanish and Dutch parameters (Sales-Ortells et al., 2015); and b) Canadian parameters (Chhipi-Shrestha et al., 2017).

A Monte Carlo simulation of 2 x 10^5 iterations was used. Probability distributions were used for most input parameters, and when distributions were fitted to available data sets, parameters were determined using maximum likelihood fitting and chi-squared goodness of fit statistics. All modeling and analyses were conducted in Mathematica 11® (Wolfram Research, 2017). For all model scenarios, 90% quantiles were calculated using the percentile method. The sensitivity analysis was performed following two complementary approaches: a. the Spearman correlation of each input parameter was determined with the daily probability of illness as the output parameter (Vose, 2008), and b. the Fourier Amplitude Sensitivity Test (FAST) which estimates the contribution of different inputs to the variance of the output (Cukier et al., 1973).

Results

The estimates of viral concentrations of HAdV and NoV GII in raw sewage and secondary and tertiary effluent by WWTPs, including the number of positive samples, are described in Table 2. In this study, the pathogen concentrations in the tertiary effluent of each WWTP and in a joint model were estimated using the pathogen concentrations characterized in the raw sewage, reduction due to secondary treatment, and the total reduction (secondary and tertiary treatments together); parameters were evaluated to determine the differences between the intake concentration and the treatment efficiency by likelihood-ratio tests (Table 3).

Table 2: Concentrations of HAdV and NoV GII (genome copies (GC)/100 ml) in each WWTP and by type of water (see supplementary materials Table S1 for complete database).

Virus (samples)	Water	WWTP 1			WWTP 2		
		+	Mean[a]	sd	+	Mean[a]	sd
HAdV (12)	Raw sewage	12	1.98×10^5	3.15×10^5	12	6.72×10^4	7.04×10^4
	Secondary	10	2.06×10^4	3.55×10^4	12	9.62×10^3	2.54×10^4
	Tertiary	9	4.3×10^2	5.66×10^2	4	7.70×10^1	2.36×10^2
NoV GII (12)	Raw sewage	12	5.17×10^6	8.88×10^6	12	2.3×10^6	3.67×10^6
	Secondary	10	3.17×10^5	8.86×10^5	9	6.32×10^4	9.11×10^4
	Tertiary	5	1.65×10^4	2.36×10^4	3	8.22×10^1	1.8×10^2

(+) Number of positive samples; (a) mean (GC/100 ml) based on the total number of samples.

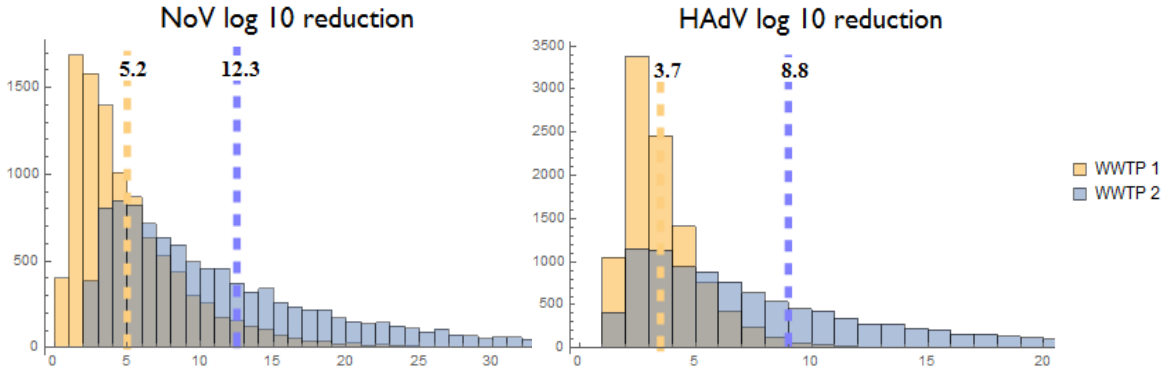
The deviances (-2*log-likelihood) in raw sewage showed that the concentration of NoV GII was the same in both WWTPs (p-value 0.408) and that the HAdV concentration was higher in WWTP 1 (p-value 0.037) than WWTP 2. Moreover, the viral concentrations after secondary treatment were also the same in both WWTPs (p-values of 0.072 and 0.287 for HAdV and NoV GII, respectively), but the concentrations after both secondary and tertiary treatments were higher in WWTP 1 for both viruses (p-values 0.009 and 0.04 for HAdV and NoV GII, respectively). This result means that the wetland removed more of both viruses than the tertiary disinfection.

Maximum likelihood estimates for the best fit of the HAdV and NoV GII concentration data described the raw concentrations and concentrations that had been reduced due to the whole treatment (Table 3). The mean $\log_{10}$ reduction of HAdV and NoV GII concentrations due to secondary treatments for the two WWTPs was 1.42 (95% confidence intervals: 0.42, 3.35) and 1.71 (0.73, 3.59), respectively. The log transformations of both viruses in the Beta distribution that describes the whole treatment efficiency were represented in Figure 1 to demonstrate the differences between the WWTPs in terms of each virus.

Table 3: Maximum likelihood gamma and beta distribution parameters fitted to reported HAdV and NoV GII concentrations (GC/100 ml) in raw samples and after full treatment from both WWTPs.

Virus	WWTP	Raw sewage parameters		Reduction from raw to tertiary treatment	
		r	λ	α	β
HAdV	1	0.92	215675	0.26	75.6
	2	1.24	54162.5	0.06	42.2
NoV GII	1	0.46	1.02×10^7	0.1	7.41
	2	0.34	5.86×10^6	0.05	373

Figure 1: Log reduction in NoV (left) and HAdV (Right) concentrations in WWTP 1 (yellow/light color) and WWTP 2 (blue/dark color). Mean values are represented with dashed lines and the respective colors of the samples.



The estimations of viral concentrations as well as the main steps of the exposure assessment, including dose and the risk estimates in terms of viruses and WWTPs, are summarized in Table 4. The limited efficiency of virus removal by tertiary treatments results in the disease burden in almost all the evaluated cases not satisfying the limit established by the WHO (10^{-6} DALYs per year).

Table 4: Mean results of QMRA and percentile 95 for the irrigation of lettuce with tertiary-treated water of two WWTPs and using HAdV and NoV GII as pathogen indicators.

			HAdV		NoV GII	
	Outputs	unit	mean	95%	mean	95%
WWTP1	Concentration after tertiary treatment	GC/ml	6.68	33.03	645.13	2829.54
	Concentration at consumption	virus/g	2.33E-05	8.64E-05	0.0592	0.1559
	dose	pppd	4.51E-04	1.15E-03	1.14	1.59
	Daily Probability of infection	pppd	2.86E-04	7.45E-04	0.0399	0.3520
	Daily probability of disease	pppd	1.45E-04	3.73E-04	0.0279	0.2470
	Yearly probability of disease	pppy	3.06E-02	7.01E-02	0.9970	0.9998
	DALYs	DALYs/year	1.44E-03	3.31E-03	1.94E-03	2.00E-03
WWTP2	Concentration after tertiary treatment	GC/ml	0.944	4.29	2.45	5.43
	Concentration at consumption	virus/g	3.27E-06	7.60E-06	2.31E-04	2.53E-04
	dose	pppd	6.27E-05	6.95E-05	5.02E-03	1.87E-03
	Daily Probability of infection	pppd	4.02E-05	4.49E-05	1.11E-03	8.25E-04
	Daily probability of disease	pppd	1.98E-05	2.30E-05	7.75E-04	5.78E-04
	Yearly probability of disease	pppy	4.23E-03	1.19E-02	0.153	0.382
	DALYs	DALYs/year	2.09E-04	5.87E-04	2.99E-04	7.47E-04

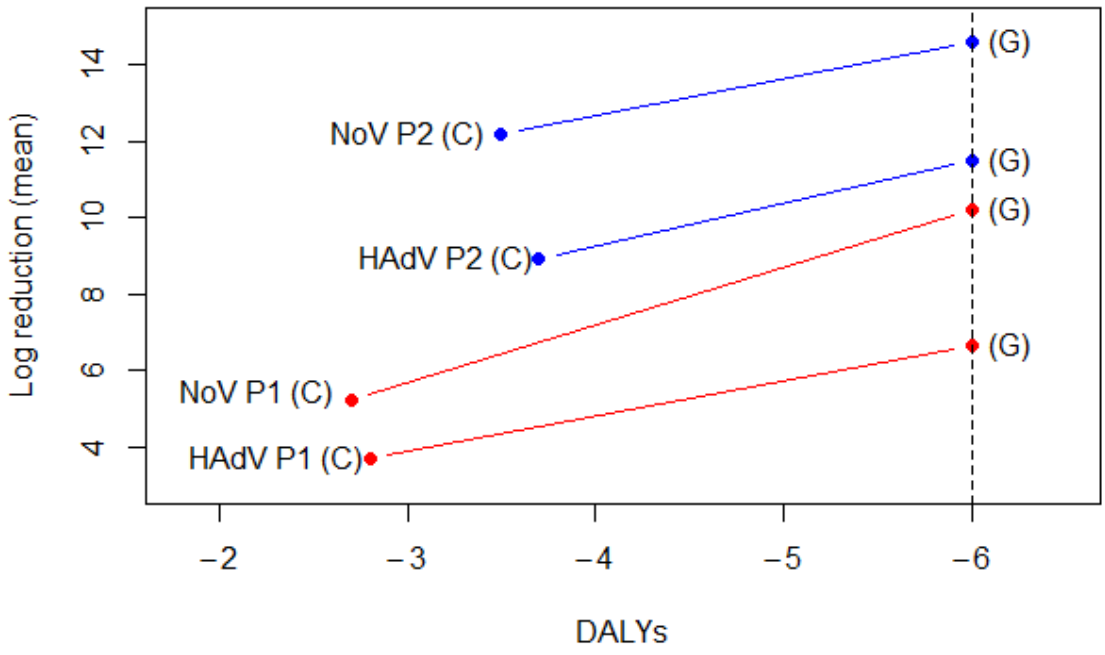
pppd: per person per day; pppy: per person per year; GC: genome copies

Sensitivity analysis using the daily probability as an output suggests that the reduction in viral concentration due to treatment, the viral concentration in raw sewage and virus ingestion were the most sensitive parameters that impact the probability of illness and burden of disease (see supplementary material Tables S3-S6 for details).

With the models fully developed, we estimated the virus concentration in tertiary effluent and the virus log reduction necessary to reach the DALYs suggested by the WHO, modifying the beta shape parameter of the Beta distribution, which describes the efficiency of each WWTP. Figure 2 shows the current and goal log reduction necessary to achieve the WHO recommendation, and additional detailed information is provided in the supplementary materials (Table S2). Although WWTP1 showed an average lower reduction in the virus concentrations, this plant required less reduction than WWTP2 to reach 10^{-6} DALYs. Additionally, we estimated a concentration that was required to reach 10^{-6} DALYs/pppy, and this value was considered the maximum level of tolerance in the

concentration of viruses in tertiary treatments. The value to reach 10^{-6} DALYs/pppy was 0.5 GC/100 ml for both viruses in the reclaimed water used for the irrigation of fresh vegetables.

Figure 2: Log reduction value plotted by virus and WWTP against the annual disease burden. Left (C) points represent the current log reduction and its corresponding disease burden for each virus and WWTP. Right (G) points are the log reduction required to achieve the WHO recommendation of 10^{-6} DALYs. Red lines represent WWTP1, and blue lines represent WWTP2.



Discussion

In this study, the concentrations of HAdV and NoV GII were quantified monthly for one year in two WWTPs and analyzed to characterize the viral concentrations in raw sewage and treated effluents. The changes in viral concentrations by two WWTPs with conventional secondary treatments but different tertiary treatments were compared. The virus concentrations found in raw sewage were on average approximately the expected values and similar to those of other raw sewage in

Mediterranean areas (Calgua et al., 2013b; Iaconelli et al., 2017) and worldwide that were evaluated with the same method of quantification, q(RT)PCR (Campos et al., 2016; Grøndahl-Rosado et al., 2014; Hata et al., 2013). Although q(RT)PCRs overestimate virus concentrations because they do not differentiate between infectious and non-infectious viral particles, which also underestimates the treatment efficacy, q(RT)PCRs are the method of choice to quantify viruses in water because they are more efficient in detecting viruses. Moreover, q(RT)PCR is currently the only method available to quantify NoV with reasonable accuracy and precision (Gerba et al., 2017).

The concentrations and concentration reductions due to secondary treatment are also in the range found for other WWTPs with active sludge treatment (Campos et al., 2016; Hata et al., 2013; Sales-Ortells et al., 2015; Sano et al., 2016). However, the reductions in virus concentration by the whole treatment were different for the WWTPs and showed more variability than expected. These results suggest that the wetland in WWTP2 was more efficient in reducing virus concentrations, although the large variability in the treatment results and the surface area required to treat the water makes this process difficult to apply to large volumes. This WWTP treated only 10% of the total volume in the wetland. Little information is available in the literature about virus removal in treatment wetlands, but lower values were found than those reported in this study, and approximately 2 logs of reduction were observed for coliphages (including somatic, F+ and MS2 coliphages); the reduction by WWTP2 was more similar to that previously observed for enterovirus (4 logs) (Barrett et al., 2001; Kadlec and Wallace, 2009). The more complex treatment in WWTP 1 (UV, chlorination and Actiflo®) has a lower reduction but a greater control of variability in the process at the expense of a greater expenditure of energy. The reduction in viral concentration by the whole treatment process in this WWTP is also in the range previously reported for similar treatment processes (Campos et al., 2016; Iaconelli et al., 2017)

Treatments with activated sludge, chlorination and sand filtration achieve approximately 3 to 5 logs of reduction in *E. coli*. However, the viral reduction with the same treatments would be between 1 and 3.5 logs, which means that WWTPs are not efficient enough to address viral reductions in water (Hata et al., 2013; Ottoson et al., 2006; Petterson and Stenström, 2015; Sano et al., 2016). Fecal coliform bacteria are much more readily inactivated by free chlorine in comparison to more persistent viruses and protozoa (Ashbolt et al., 2001). Other known tertiary treatments, such as Actiflo®, are recognized to reduce coliphage loads between 1 and 3 logs under experimental conditions, but the reduction depends on several factors such as the wastewater quality and sensitivity of the target microorganisms to the treatment (Mok et al., 2014).

Adequate characterization of pathogen concentrations is essential for making appropriate risk assessments. Mathematical models have thus been developed to address this problem and produce a better approach by considering viral concentrations before and after treatment (Teunis et al., 1999, 2009). Microbial monitoring before and after treatment is the most direct way to assess treatment efficacy (Smeets et al., 2010), and these methods have been recommended in QMRA analysis (WHO, 2016). The input and output samples were considered unpaired because sampling a body of water in exactly the same way before and after the treatment process is complicated. The viral quantifications were used to establish a distribution of values that described the concentrations of viruses in raw sewage and the treatment efficacy for viruses by WWTPs. These distributions allow the inclusion of particularly extreme values that are sometimes found in outbreak situations. These distributions were included within a QMRA framework recommended by WHO for the irrigation of vegetables with reclaimed water (WHO, 2016, 2006).

Most of the other parameters used in exposure assessment for the irrigation of vegetables in this study have been described in previous risk assessment studies published worldwide. However, our study includes modifications that we consider important to describe the correct dose. One

modification was including the recovery of SMF concentration method (Gonzales-Gustavson et al., 2017). Virus recovery rates from concentrating and molecular methods can be quite low, resulting in underestimations of the real concentration by 1 to 3 orders of magnitude (Mok and Hamilton, 2014; Petterson et al., 2015). The recovery efficiency of a model needs to account for each method of concentration. The advantage to stochastically including recovery is that these values vary between samples. However, the recovery with SMF concentrated viruses in water with low variability (approximately 69% for HAdV and 40% for NoV) (Gonzales-Gustavson et al., 2017)

Another main component of QMRA is the dose-response model, which describes a relation between the dose and the likelihood of infection or illness outcomes. For this reason, the choice of a dose-response model can be highly relevant in the overall determination of risk. Several dose-response models are available (Van Abel et al., 2017). Although some publications used the Beta-Poisson approximation and an exponential for NoV and HAdV, respectively, we chose the recently published HAdV model, which has the advantage of been stablished specifically for oral inoculation (Teunis et al., 2016) in contrast with the exponential dose-response model used for inhalation. The latter method was based on a respiratory HAdV strain, which limited its use in QMRA studies (Ashbolt, 2015). In addition, the hypergeometric dose-response function for NoV may include the effects of viruses that are aggregated or not, which is important because in environmental samples, this virus may be in different aggregation states. The Beta-Poisson model might not accurately approximate the dose-response function when little information is available (Teunis et al., 2008; Teunis and Havelaar, 2000). The assumption that aggregation occurs is less applicable for treated water since treatment processes remove large particles more effectively than small particles. Additionally, models that include the effects of aggregation tends to yield a lower probability of infection than models that do not include it, particularly at lower doses, thereby underestimating the risk (Mcbride, 2014; Van Abel et al., 2017).

To estimate DALYs, we used parameters described previously by Sales-Ortells et al. (Sales-Ortells et al., 2015) as the years lived with disability plus the years of life lost; these values describe a mix of values from Catalonia and the Netherlands due to a lack of available information (Sales-Ortells et al., 2015). The results were similar to the parameters of disease burden per case and susceptibility fraction described in research from Canada (Chhipi-Shrestha et al., 2017) (data not shown). We used the Canadian parameters for the estimation of DALYs for this virus because no disease burden parameters for HAdV in Catalonia are available. Although immunity to NoV infections is not relevant to modifying the proportion of susceptible individuals and the proportion of secretor-negative members of the Hispanic population is negligible (approximately 2%) (Van Abel et al., 2017), the situation for HAdV would be different, and either immunity or the spatial distribution of individuals susceptible or resistant to HAdV infection would affect our model. Future models that include these parameters are necessary to better estimate DALYs for HAdV in the studied region.

The QMRA results demonstrate that the systems fail to achieve the actual recommendation by the WHO of 10^{-6} DALYs pppy in both WWTPs and with both viruses. Both WWTPs therefore failed to meet the threshold for acceptable risk levels, indicating that the virus removal capacities of these treatments were insufficient and that other disinfection or barrier treatments must be considered before the irrigation of lettuce. Both WWTPs require additional reduction to achieve the WHO goal. However, due to the high variability in the treatment of WWTP 2, the goal is more likely to be achieved in WWTP 1. The quantification of tertiary treatment effectiveness in WWTP2 showed lower viral loads, with 9 of 12 negative samples detected through the year of evaluation. For that reason, the simulated distributions of the reduction due to treatment showed longer right tails and higher mean reduction than found in WWTP 1.

One of the main problems in QMRA studies is the lack of information available to establish a distribution to describe the concentration of microorganisms and the reduction in the WWTPs.

Unfortunately, methods to quantify viruses after treatment often yield negative results or values that are below the limits of quantification because of their low sensitivity and the need of testing high volumes of water for accuracy. Negative results do not mean that viruses were completely removed, as the concentration of the FIB suggests (De Keuckelaere et al., 2015; Mok et al., 2014; Petterson and Ashbolt, 2016; Schijven et al., 2011; WHO, 2016).

Since sewage and secondary and tertiary effluents are not routinely tested for viruses, the occurrence of human enteric viruses in water remains largely unknown unless an outbreak is reported, and the samples that are usually collected seldom demonstrate the viral origin (Gibson, 2014; Gorchev and Ozolins, 2011). Unfortunately, the limits of detecting HAdV and NoV GII using the SMF method are 28.6 and 291 GC/100 ml, making use of a single evaluation to determine whether WWTPs are efficient enough to satisfy the WHO demands impossible (Table 5 and Figure 4). However, our study shows that these assays can be used in field evaluations of the concentrations of HAdV and NoV GII in sewage and of the removal efficacy of secondary and tertiary treatment processes, thus providing a foundation of evidence to assess the safety of reclaimed water systems for food crop irrigation and for the required virus removal to provide water safe for unrestricted irrigation.

The health risk associated with the consumption of lettuce irrigated with tertiary-treated effluent from two WWTPs, considering NoV GII and HAdV, has been estimated based on the quantification of realistic viral loads in the treatment. The results suggest that HAdV could be used as reference pathogen to validate WWTP treatments as it shows similar risk values as NoV GII.

Conclusions

To assess the health risk associated with reclaimed water, we used a stochastic QMRA model to estimate the annual disease burden from the consumption of lettuce irrigated with tertiary-treated water from two different WWTPs. Major findings were

- None of the WWTPs, on average, met the threshold of $\leq 10^{-6}$ DALY pppy for an acceptable level of risk.
- The results from the WWTPs that applied wetland and conventional tertiary treatments differed, with the wetland treatment giving better reductions (12.2 and 8.9 logs for NoV GII and HAdV, respectively) than the conventional treatment (5.2 and 3.7 logs) but with more variation than a conventional treatment with UV, chlorination and Actiflo®.
- This report is the first description of a QMRA assay developed with HAdV in regard to the irrigation of vegetables, and this assay showed similar results to the assays with NoV GII, even in the wetland-treated samples.
- Reduction requirements in tertiary effluents, to reach the WHO recommended levels, depend on the treatments developed in the WWTPs, and additional average reductions are necessary (in WWTP 1: 3 logs for HAdV and 4.8 logs for NoV GII; and in WWTP 2: 2.6 logs for HAdV and 2.4 logs for NoV GII).

Acknowledgements

The study reported here was partially funded by the projects JPI Water METAWATER (4193-00001B), AGL2014-55081-R from Spanish MINECO, and the VIROCLIME, funded under the EU Seventh Framework Program, Contract No. 243923. During the development of this study, Eloy Gonzales-Gustavson was a fellow of the Peruvian Government. We acknowledge the Ajuntament de Granollers, and especially X. Hidalgo, for their helpful collaboration.

References

Adefisoye, M.A., Nwodo, U.U., Green, E., Okoh, A.I., 2016. Quantitative PCR Detection and Characterisation of Human Adenovirus, Rotavirus and Hepatitis A Virus in Discharged Effluents of Two Wastewater Treatment Facilities in the Eastern Cape, South Africa. Food Environ. Virol. 8, 262–274. https://doi.org/10.1007/s12560-016-9246-4

Aecosan, 2015. Encuesta ENALIA 2. Encuesta Nacional de Alimentación en población adulta, mayores y embarazadas. Agencia española Consum. Segur. y Nutr.

Albinana-Gimenez, N., Clemente-Casares, P., Calgua, B., Huguet, J.M., Courtois, S., Girones, R., 2009. Comparison of methods for concentrating human adenoviruses, polyomavirus JC and noroviruses in source waters and drinking water using quantitative PCR. J. Virol. Methods 158, 104–109. https://doi.org/10.1016/j.jviromet.2009.02.004

Allende, A., Monaghan, J., 2015. Irrigation water quality for leafy crops: A perspective of risks and potential solutions. Int. J. Environ. Res. Public Health 12, 7457–7477. https://doi.org/10.3390/ijerph120707457

Ashbolt, N., Grabow, W., Snozzi, M., 2001. Indicators of microbial water quality, in: Fewtrell, L., Bartram, J. (Eds.), Water Quality: Guidelines, Standards and Health. IWA Publishing, London, pp. 289–316. https://doi.org/10.4324/9781315693606

Ashbolt, N.J., 2015. Microbial Contamination of Drinking Water and Human Health from Community Water Systems. Curr. Environ. Heal. reports 2, 95–106. https://doi.org/10.1007/s40572-014-0037-5

Atmar, R.L., Opekun, A.R., Gilger, M.A., Estes, M.K., Crawford, S.E., Neill, F.H., Ramani, S., Hill, H., Ferreira, J., Graham, D.Y., 2014. Determination of the 50% human infectious dose for norwalk

virus. J. Infect. Dis. 209, 1016–1022. https://doi.org/10.1093/infdis/jit620

Barker, S.F., 2014. Risk of norovirus gastroenteritis from consumption of vegetables irrigated with highly treated municipal wastewater-evaluation of methods to estimate sewage quality. Risk Anal. 34, 803–817. https://doi.org/10.1111/risa.12138

Barrett, E.C., Sobsey, M.D., House, C.H., White, K.D., 2001. Microbial indicator removal in onsite constructed wetlands for wastewater treatment in the southeastern U.S. Water Sci. Technol. 44, 177–182.

Boletin Oficial del Estado, 2007. Real Decreto 1620/2007, de 7 de diciembre, por el que se establece el régimen jurídico de la reutilización de las aguas depuradas. Nº 294 de 8 diciembre.

Calgua, B., Fumian, T., Rusinol, 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. https://doi.org/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., Lopez-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. https://doi.org/10.1016/j.jviromet.2008.08.003

Calgua, B., Rodriguez-Manzano, J., Hundesa, A., Sunen, E., Calvo, M., Bofill-Mas, S., Girones, R., 2013b. New methods for the concentration of viruses from urban sewage using quantitative PCR. J. Virol. Methods 187, 215–221. https://doi.org/10.1016/j.jviromet.2012.10.012

Campos, C.J.A., Avant, J., Lowther, J., Till, D., Lees, D.N., 2016. Human norovirus in untreated sewage and effluents from primary, secondary and tertiary treatment processes. Water Res. 103, 224–

232. https://doi.org/10.1016/j.watres.2016.07.045

Canada, H., 2010. Canadian Guidelines for Domestic Reclaimed Water for Use in Toilet and Urinal Flushing, Ministry of Health. Ottawa ON.

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, 128–134. https://doi.org/10.1016/j.ijfoodmicro.2013.04.007

Chhipi-Shrestha, G., Hewage, K., Sadiq, R., 2017. Microbial quality of reclaimed water for urban reuses: Probabilistic risk-based investigation and recommendations. Sci. Total Environ. 576, 738–751. https://doi.org/10.1016/j.scitotenv.2016.10.105

Cukier, R.I., Fortuin, C.M., Shuler, K.E., Petschek, A.G., Schaibly, J.H., 1973. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. J. Chem. Phys. 59, 3873–3878. https://doi.org/10.1063/1.1680571

De Keuckelaere, A., Jacxsens, L., Amoah, P., Medema, G., Mcclure, P., Jaykus, L.A., Uyttendaele, M., 2015. Zero Risk Does Not Exist: Lessons Learned from Microbial Risk Assessment Related to Use of Water and Safety of Fresh Produce. Compr. Rev. Food Sci. Food Saf. 14, 387–410. https://doi.org/10.1111/1541-4337.12140

EU, 2016. EU-level instruments on water reuse: Final report to support the commission ' s impact assessment. https://doi.org/10.2779/974903

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. https://doi.org/10.1016/j.watres.2016.11.020

Ghebremedhin, B., 2014. Human adenovirus: Viral pathogen with increasing importance. Eur. J. Microbiol. Immunol. 4, 26–33. https://doi.org/10.1556/EuJMI.4.2014.1.2

Gibson, K.E., 2014. Viral pathogens in water: Occurrence, public health impact, and available control strategies. Curr. Opin. Virol. 4, 50–57. https://doi.org/10.1016/j.coviro.2013.12.005

Gonzales-Gustavson, E., Cárdenas-Youngs, Y., Calvo, M., da Silva, M.F.M., Hundesa, A., Amorós, I., Moreno, Y., Moreno-Mesonero, L., Rosell, R., Ganges, L., Araujo, R., Girones, R., 2017. Characterization of the efficiency and uncertainty of skimmed milk flocculation for the simultaneous concentration and quantification of water-borne viruses, bacteria and protozoa. J. Microbiol. Methods 134, 46–53. https://doi.org/10.1016/j.mimet.2017.01.006

Gorchev, H.G., Ozolins, G., 2011. WHO guidelines for drinking-water quality. WHO Chron. 38, 104–108. https://doi.org/10.1016/S1462-0758(00)00006-6

Grøndahl-Rosado, R.C., Yarovitsyna, E., Trettenes, E., Myrmel, M., Robertson, L.J., 2014. A One Year Study on the Concentrations of Norovirus and Enteric Adenoviruses in Wastewater and A Surface Drinking Water Source in Norway. Food Environ. Virol. 6, 232–245. https://doi.org/10.1007/s12560-014-9161-5

Haas, C.N., Rose, J.B., Gerba, C., Regli, S., 1993. Risk assessment of virus in drinking water. Risk Anal. 13, 545–552.

Hata, A., Kitajima, M., Katayama, H., 2013. Occurrence and reduction of human viruses, F-specific RNA coliphage genogroups and microbial indicators at a full-scale wastewater treatment plant in Japan. J. Appl. Microbiol. 114, 545–554. https://doi.org/10.1111/jam.12051

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. https://doi.org/10.1128/AEM.68.9.4523

Hijnen, W.A.M., Beerendonk, E.F., Medema, G.J., 2006. Inactivation credit of UV radiation for viruses, bacteria and protozoan (oo)cysts in water: a review. Water Res. 40, 3–22. https://doi.org/10.1016/j.watres.2005.10.030

Iaconelli, M., Muscillo, M., Della Libera, S., Fratini, M., Meucci, L., De Ceglia, M., Giacosa, D., La Rosa, G., 2017. One-year Surveillance of Human Enteric Viruses in Raw and Treated Wastewaters, Downstream River Waters, and Drinking Waters. Food Environ. Virol. 9, 79–88. https://doi.org/10.1007/s12560-016-9263-3

Iglesias, R., Ortega, E., Batanero, G., Quintas, L., 2010. Water reuse in spain: Data overview and costs estimation of suitable treatment trains. Desalination 263, 1–10. https://doi.org/10.1016/j.desal.2010.06.038

Kadlec, R.H., Wallace, S.D., 2009. Treatment Wetlands, Second edi. ed. CRC Press, Taylor & Francis Group, London. https://doi.org/10.1201/9781420012514

Kageyama, T., Kojima, S., Shinohara, M., Uchida, K., Fukushi, S., Hoshino, F.B., Takeda, N., Katayama, K., 2003. Broadly reactive and highly sensitive assay for Norwalk-like viruses based on real-time quantitative reverse transcription-PCR. J Clin Microbiol 41, 1548–1557. https://doi.org/10.1128/JCM.41.4.1548

Karavarsamis, N., Hamilton, A.J., 2010. Estimators of annual probability of infection for quantitative microbial risk assessment. J. Water Health 8, 365–373. https://doi.org/10.2166/wh.2010.045

Kobayashi, M., Matsushima, Y., Motoya, T., Sakon, N., Shigemoto, N., Okamoto-Nakagawa, R., Nishimura, K., Yamashita, Y., Kuroda, M., Saruki, N., Ryo, A., Saraya, T., Morita, Y., Shirabe, K.,

Ishikawa, M., Takahashi, T., Shinomiya, H., Okabe, N., Nagasawa, K., Suzuki, Y., Katayama, K., Kimura, H., 2016. Molecular evolution of the capsid gene in human norovirus genogroup II. Sci. Rep. 6, 1–11. https://doi.org/10.1038/srep29400

Kundu, A., McBride, G., Wuertz, S., 2013. Adenovirus-associated health risks for recreational activities in a multi-use coastal watershed based on site-specific quantitative microbial risk assessment. Water Res. 47, 6309–6325. https://doi.org/10.1016/j.watres.2013.08.002

Mara, D., Sleigh, A., 2010. Estimation of norovirus infection risks to consumers of wastewater-irrigated food crops eaten raw. J. Water Health 8, 39–43. https://doi.org/10.2166/wh.2009.140

Mcbride, G.B., 2014. Norovirus dose-response in sewage-related QMRA: The importance of virus aggregation. Congr. Env. Model. Softw. Paper 69.

McMinn, B.R., Ashbolt, N.J., Korajkic, A., 2017. Bacteriophages as indicators of faecal pollution and enteric virus removal. Lett. Appl. Microbiol. 65, 11–26. https://doi.org/10.1111/lam.12736

Mok, H.-F., Hamilton, A.J., 2014. Exposure Factors for Wastewater-Irrigated Asian Vegetables and a Probabilistic Rotavirus Disease Burden Model for Their Consumption. Risk Anal. 34, 602–613. https://doi.org/10.1111/risa.12178

Mok, H.F., Barker, S.F., Hamilton, A.J., 2014. A probabilistic quantitative microbial risk assessment model of norovirus disease burden from wastewater irrigation of vegetables in Shepparton, Australia. Water Res. 54, 347–362. https://doi.org/10.1016/j.watres.2014.01.060

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, 1449–1457.

https://doi.org/10.1016/j.watres.2006.01.039

Owusu-Ansah, E. de G.J., Sampson, A., Amponsah, S.K., Abaidoo, R.C., Dalsgaard, A., Hald, T., 2017. Probabilistic quantitative microbial risk assessment model of norovirus from wastewater irrigated vegetables in Ghana using genome copies and fecal indicator ratio conversion for estimating exposure dose. Sci. Total Environ. 601–602, 1712–1719. https://doi.org/10.1016/j.scitotenv.2017.05.168

Petterson, S., Grøndahl-Rosado, R., Nilsen, V., Myrmel, M., Robertson, L.J., 2015. Variability in the recovery of a virus concentration procedure in water: Implications for QMRA. Water Res. 87, 79–86. https://doi.org/10.1016/j.watres.2015.09.006

Petterson, S.R., Ashbolt, N.J., 2016. QMRA and water safety management: Review of application in drinking water systems. J. Water Health 14, 571–589. https://doi.org/10.2166/wh.2016.262

Petterson, S.R., Ashbolt, N.J., Sharma, A., 2001. Microbial risks from wastewater irrigation of salad crops: A screening-level risk assessment. Water Environ. Res. 73, 667–672. https://doi.org/10.2175/106143001x143402

Petterson, S.R., Stenström, T.A., 2015. Quantification of pathogen inactivation efficacy by free chlorine disinfection of drinking water for QMRA. J. Water Health 13, 625–644. https://doi.org/10.2166/wh.2015.193

Pina, S., Puig, M., Lucena, F., Jofre, J., Girones, R., 1998. Viral pollution in the environment and in shellfish: Human adenovirus detection by PCR as an index of human viruses. Appl. Environ. Microbiol. 64, 3376–3382.

Rames, E., Roiko, A., Stratton, H., Macdonald, J., 2016. Technical aspects of using human adenovirus as a viral water quality indicator. Water Res. 96, 308–326.

https://doi.org/10.1016/j.watres.2016.03.042

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. https://doi.org/10.1016/j.watres.2014.04.013

Rusiñol, M., Fernandez-Cassi, X., Timoneda, N., Carratala, 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. https://doi.org/10.1016/j.jenvman.2015.05.019

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. https://doi.org/10.1016/j.foodres.2014.08.018

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. https://doi.org/10.1016/j.envint.2016.03.001

Sanz, L.A., Gawlik, B.M., 2014. Water Reuse in Europe, Relevant guidelines, needs for and barriers to innovation Third, E-WAter (European Water Association). https://doi.org/10.2788/29234

Schijven, J.F., Teunis, P.F.M., Rutjes, S.A., Bouwknegt, M., de Roda Husman, A.M., 2011. QMRAspot: A tool for Quantitative Microbial Risk Assessment from surface water to potable water. Water Res. 45, 5564–5576. https://doi.org/10.1016/j.watres.2011.08.024

Smeets, P.W.M.H., Rietveld, L.C., Van Dijk, J.C., Medema, G.J., 2010. Practical applications of quantitative microbial risk assessment (QMRA) for water safety plans. Water Sci. Technol. 61, 1561–1568. https://doi.org/10.2166/wst.2010.839

Teunis, P., Evers, E., Slob, W., 1999. Analysis of variable fractions resulting from microbial counts. Quant. Microbiol. 1, 63–88. https://doi.org/10.1023/A:1010028411716

Teunis, P., Moe, C.L., Liu, P., Miller, S., Lindesmith, L., Baric, B., Pendu, J., Calderon, R., 2008. Norwalk Virus: How infectious is it? Anticancer Res. 80, 1468–1476. https://doi.org/10.1002/jmv.21237

Teunis, P., Schijven, J., Rutjes, S., 2016. A generalized dose-response relationship for adenovirus infection and illness by exposure pathway. Epidemiol. Infect. 144, 3461–3473. https://doi.org/10.1017/S0950268816001862

Teunis, P.F.M., Havelaar, A.H., 2000. The Beta Poisson Dose-Response Model Is Not a Single-Hit Model 20.

Teunis, P.F.M., Rutjes, S.A., Westrell, T., de Roda Husman, A.M., 2009. Characterization of drinking water treatment for virus risk assessment. Water Res. 43, 395–404. https://doi.org/10.1016/j.watres.2008.10.049

Van Abel, N., Schoen, M.E., Kissel, J.C., Meschke, J.S., 2017. Comparison of Risk Predicted by Multiple Norovirus Dose-Response Models and Implications for Quantitative Microbial Risk Assessment. Risk Anal. 37, 245–264. https://doi.org/10.1111/risa.12616

Vose, D., 2008. Risk analysis a quantitative guide, third edit. ed. John Wiley and Sons, Wiltshire, England.

WHO, 2016. Quantitative Microbial Risk Assessment: Application for Water Safety Management. Geneva, Switzerland. https://doi.org/10.1002/9781118910030

WHO, 2006. WHO guidelines for the safe use of wastewater , excreta and greywater: Wastewater use in agriculture II, 204. https://doi.org/10.1007/s13398-014-0173-7.2

Wolfram Research, I., 2017. Mathematica, Versión 11. ed, Wolfram Research, Inc. Wolfram Research, Inc., Champaign, Illinois.

Wyn-Jones, A.P., Carducci, A., Cook, N., D'Agostino, M., Divizia, M., Fleischer, J., Gantzer, C., Gawler, A., Girones, R., Höller, C., de Roda Husman, A.M., Kay, D., Kozyra, I., López-Pila, J., Muscillo, M., José Nascimento, M.S., Papageorgiou, G., Rutjes, S., Sellwood, J., Szewzyk, R., Wyer, M., 2011. Surveillance of adenoviruses and noroviruses in European recreational waters. Water Res. 45, 1025–1038. https://doi.org/10.1016/j.watres.2010.10.015

4.3 Artículo 3: Identificación de sapovirus GV.2, astrovirus VA3 y nuevos anellovirus in sueros de pacientes con hepatitis agudo de etiología desconocida

Identification of sapovirus GV.2, astrovirus VA3 and novel anelloviruses in serum from patients with acute hepatitis of unknown aetiology

Eloy A. Gonzales Gustavson, N. Timoneda, X. Fernandez-Cassi, A. Caballero, J. F. Abril, M. Buti, F. Rodriguez-Frias, R. Girones

El termino hepatitis se define como la inflamación del hígado, la cual es causada principalmente por virus, aunque también puede ser debida a muchos otros diversos factores. A pesar del gran avance e inversión en el estudio de la hepatitis, entre el 4% y el 34% de los casos no se puede llegar a diagnosticar el origen del problema. Los virus conocidos como causantes de hepatitis aguda tienen transmisión fecal-oral, por lo que pueden utilizar el agua contaminada como vía para llegar a la población susceptible. En este estudio se plantea el objetivo de identificar virus de posible transmisión hídrica que podrían haber causado hepatitis aguda de etiología desconocida. Con este fin se analizó suero de pacientes con signos clínicos de hepatitis aguda mediante técnicas de metagenómica para caracterizar la población viral de estos pacientes. Por motivos prácticos, las muestras fueron agrupadas en pools. Se hicieron cuatro pools de pacientes con hepatitis de etiología desconocida: hombres menores de 45 años (Male A pool), hombres mayores de 45 años (Male B pool), mujeres (Female pool) y pacientes autoinmunes/inmunosuprimidos (Ai+ImSP pool). Adicionalmente, se incluyeron: un pool de pacientes con hepatitis E (HEV pool) y pools de voluntarios sanos como controles positivo y negativo respectivamente. Se encontraron una gran variedad de virus: en primer lugar, los anellovirus, dentro de ellos algunas secuencias que no habían sido descritas previamente y en mayor frecuencia en los pools conformados de pacientes con etiología desconocida. Adicionalmente se encontraron virus recientemente descritos y que han sido asociados a gastroenteritis como sapovirus GV.2 y astrovirus VA3, los cuales se detectaron exclusivamente en los pools de pacientes con hepatitis de etiología desconocida. También se logró amplificar el 76% del genoma completo del HEV en el pool considerado como

control positivo. Finalmente, otros virus como el GB virus C y retrovirus endógenos humanos se detectaron tanto en pools con hepatitis como en los pools de voluntarios sanos. Este estudio provee una visión general del viroma del suero de pacientes con hepatitis sugiriendo la presencia de virus no descritos antes en casos de hepatitis. Todas las familias virales identificadas como potenciales agentes de hepatitis serían capaces de ser transmisibles a través del agua y podrían implicar un riesgo a la población a través de esta vía. Sin embargo, es necesario realizar futuros estudios epidemiológicos que confirmen su asociación con el desarrollo de la hepatitis.

PLOS ONE

RESEARCH ARTICLE

Identification of sapovirus GV.2, astrovirus VA3 and novel anelloviruses in serum from patients with acute hepatitis of unknown aetiology

Eloy Gonzales-Gustavson[1☯], N. Timoneda[1,2,3☯], X. Fernandez-Cassi[1], A. Caballero[4], J. F. Abril[2,3], M. Buti[4], F. Rodriguez-Frias[4], R. Girones[1]*

1 Laboratory of Virus Contaminants of Water and Food, Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, Barcelona, Catalonia, Spain, **2** Computational Genomics Lab, Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, Barcelona, Catalonia, Spain, **3** Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Catalonia, Spain, **4** Hospital Universitari Vall d'Hebron and CIBEREHD del Instituto Carlos III, Barcelona, Catalonia, Spain

☯ These authors contributed equally to this work.
* rgirones@ub.edu

Citation: Gonzales-Gustavson E, Timoneda N, Fernandez-Cassi X, Caballero A, Abril JF, Buti M, et al. (2017) Identification of sapovirus GV.2, astrovirus VA3 and novel anelloviruses in serum from patients with acute hepatitis of unknown aetiology. PLoS ONE 12(10): e0185911. https://doi.org/10.1371/journal.pone.0185911

Editor: Sibnarayan Datta, Defence Research Laboratory, INDIA

Received: May 3, 2017

Accepted: September 21, 2017

Published: October 5, 2017

Data Availability Statement: The raw sequencing data used to perform this analysis along with the FASTQ file are located in the NCBI Sequence Read Archive; BioProject (PRJNA379441).

Funding: The study reported here was partially funded by the Programa RecerCaixa 2012 (ACUP-00300), AGL2011-30461-C02-01/ALI and AGL2014-55081-R from Spanish MINECO. This study was partially funded by a grant from the

Abstract

Hepatitis is a general term meaning inflammation of the liver, which can be caused by a variety of viruses. However, a substantial number of cases remain with unknown aetiology. We analysed the serum of patients with clinical signs of hepatitis using a metagenomics approach to characterize their viral species composition. Four pools of patients with hepatitis without identified aetiological agents were evaluated. Additionally, one pool of patients with hepatitis E (HEV) and pools of healthy volunteers were included as controls. A high diversity of anelloviruses, including novel sequences, was found in pools from patients with hepatitis of unknown aetiology. Moreover, viruses recently associated with gastroenteritis as sapovirus GV.2 and astrovirus VA3 were also detected only in those pools. Besides, most of the HEV genome was recovered from the HEV pool. Finally, GB virus C and human endogenous retrovirus were found in the HEV and healthy pools. Our study provides an overview of the virome in serum from hepatitis patients suggesting a potential role of these viruses not previously described in cases of hepatitis. However, further epidemiologic studies are necessary to confirm their contribution to the development of hepatitis.

Introduction

Hepatitis is a general term meaning inflammation of the liver and can be caused by a variety of viruses, such as hepatitis A, B, C, D and E [1]. Infectious agents such as bacteria, fungi or parasites, as well as non-infectious agents such as alcohol, drugs or autoimmune diseases, may cause hepatitis too. According to the estimates of the Global Burden of Disease study, viral hepatitis is responsible for approximately 1.5 million deaths each year, which is comparable to

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, Eloy Gonzales-Gustavson is a fellow of the Peruvian Government; Natalia Timoneda is a fellow of the Spanish Ministry of Science and Xavier Fernandez-Cassi was a fellow of the Catalan Government "AGAUR" (FI-DGR).

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

the number of annual deaths from HIV/AIDS (1.3 million), malaria and tuberculosis (TB) (0.9 million and 1.3 million, respectively) [2].

Viral hepatitis is still one of the key causes of acute liver failure (ALF) in the world. ALF is a devastating clinical syndrome associated with high mortality in the absence of immediate care, specific treatment or liver transplantation [3]. Globally, hepatitis A, B and E infections are probably responsible for the majority of ALF cases. However, despite significant progress in the diagnosis and treatment of hepatitis, in a considerable number of patients, the aetiological agents remain unknown. Previous studies have found that between 3.8% and 33.9% of hospital inpatients with acute hepatitis had non-A-E-hepatitis [4–8]. Additionally, 10% of patients with ALF had non-A-E hepatitis [9].

Therapeutic trials using interferon-α to treat hepatitis of unknown aetiology have consistently resulted in response rates of approximately 50%, indicating a virological aetiology [10]. This evidence suggests that other viruses may be responsible for hepatitis. As a result, new viruses, including a *Flaviviridae* GB virus type C (GBV-C) [11] and *Anelloviridae* TTV and SEN virus [12], have been reported in recent years to be associated with hepatitis. However, epidemiological data failed to confirm a causative role for those viruses in the development of hepatitis, and a high percentage of individuals infected by them were found to be healthy carriers [13,14]. Recent investigation has shown that other viral infections such as cytomegalovirus and Epstein Barr virus may mimic viral hepatitis [15]. Less frequently, hepatitis may be present in people with herpes simplex virus [16], parvovirus B19 [17], and adenoviruses 1, 2, 5, 12 and 32 [18,19].

Epidemiologic information related to non-A-E hepatitis is scarce. In a study by Delic et al. 2010, analysing 408 patients with acute hepatitis, history of blood transfusion, drug use or other parenteral exposure were not associated with the onset of illness [7], suggesting that if the viral nature of non-A-E hepatitis is proven, it should spread primarily by non-parenteral means. Moreover, some patients diagnosed with acute non-A-E hepatitis show biochemical features at admission similar to those associated with other viral hepatitis. Apparently, acute non-A-E hepatitis is distributed worldwide, and progression to chronicity was observed in approximately 9% of patients [7,20].

The cause of acute non-A-E hepatitis remains unknown. It seems likely that another as-yet-unidentified infectious agent(s) exists [20]. Recent rapid progress in sequencing technologies and associated bioinformatics methodologies has enabled a more in-depth view of the structure and functioning of viral communities, supporting the characterization of emerging viruses [21]. With the advent of metagenomics studies, our knowledge of the different components and the complexity of the microbiome greatly expanded. The eukaryotic virome comprises viruses infecting the host, endogenous viral elements, and viruses associated with other eukaryotic components of the ingesta [22].

In this study, next-generation sequencing (NGS) was used to identify viruses in serum samples from patients suffering from acute hepatitis signs. For that purpose, the viromes in the serum of patients with Non-A-E hepatitis were analysed and the results were compared with the viromes from patients with acute hepatitis E (positive controls) and healthy patients (negative controls).

Materials and methods

Serum samples

A total number of 42 serum samples were collected from patients with acute viral hepatitis from the Vall d'Hebron Hospital, Barcelona, Spain. The clinical diagnosis of acute viral hepatitis was based on the lack of previous history of chronic liver disease, a rise in serum

aminotransferase (AST, ALT) activity of at least 200 IU/L, high values of total (TB) and direct bilirubin (DB) and exclusion of other causes of liver disease such as hepatitis A (Ig-M negative), hepatitis B (surface-antigen-HBsAg- and anti-core antibodies-anti-HBc-negative-), hepatitis C (anti-VHC-negative) and hepatitis E (HEV) (IgG, IgM and RT-PCR, all negatives). Of the 32 patients with acute hepatitis of unknown aetiology, 19 were male, and 13 were female, with ages ranging from one to 92 years old. Eight of those patients were diagnosed with an autoimmune or immunosuppressed (Ai+ImSP) condition. Additionally, serum samples from 10 patients—positive for HEV by nested RT-PCR—were included as positive controls. In addition, serum samples from 20 healthy volunteers were also evaluated.

The serum samples were pooled according to the following criteria. Patients with acute hepatitis were grouped into five pools: male pool A (8 samples, age range from 1 to 44), male pool B (8 samples, age range from 45 to 78), a female pool (8 samples, age range from 6 to 92), and an Ai+ImSP pool (8 samples, age range from 2 to 84) that included patients with the Ai+ImSP condition. Finally, a pool of HEV RNA-positive patients (10 samples, age range from 6 to 84) was included. Healthy volunteers' serum samples were grouped in two pools and evaluated in duplicate: Healthy A1 and A2 pools, with 10 females (age range between 27 and 63), and Healthy B1 and B2 pools, with 2 males and 8 females (age range between 26 and 58).

Sample preparation

Serum samples were kept at -80˚C prior to the metagenomics analysis protocol. Pools were prepared with the corresponding serum samples to achieve an initial volume of 500 μL. Briefly, the pools were first filtered through a pore size of 0.45 μm (Millipore Corp., Billerica, MA, USA) to remove cellular debris, ultracentrifuged at 100,000 × g for 90 min at 4˚C and re-suspended in 500 μL of PBS 1X. Next, 300 μL of the re-suspended pool was subjected to DNAse treatment to eliminate background DNA with 20 U TURBO™ DNase (Ambion, Thermo Fisher Scientific, Waltham, MA, USA). Then, viral nucleic acids (NAs) were extracted with QIAmp Viral RNA Mini Kit (Qiagen, Inc., Valencia, CA), without carrier RNA, according to the manufacturer's instructions. To enable the detection of both DNA and RNA viruses, total NAs were reverse-transcribed as previously described [23,24]. In short, SuperScript II (Life Technologies, California, USA) was used to retro-transcribe 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). PCR amplification with AmpliTaqGold (Life Technologies, Austin, Texas, USA) was performed using primerB (`5′-GTTTCCCAGTCACGATC-3′`) with 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 kit (cat n˚ D4013, Zymo Research, USA) to yield enough DNA for the library preparation.

Sequencing protocol

NGS sequencing was performed at SGB-UAB, Barcelona. dsDNA samples were quantified by Qubit 2.0 (Life technologies), and libraries were constructed using a Nextera XT DNA sample preparation kit (Illumina Inc). Samples were sequenced on Illumina MiSeq 2x300; all samples were multiplexed and distributed within three independent sequencing runs.

NGS data processing

The quality of raw and clean read sequences was assessed using FASTX-Toolkit software, version 0.0.14 (Hannon Lab) [25]. The sequenced reads were cleaned by Trimmomatic version 0.32 [26] while the sequencing adaptors and linker contamination were removed. Low-quality

ends were trimmed using a Phred score average threshold above Q15 over a running window of four nucleotides. Low-complexity sequences, mostly 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 and the sequence string-compression ratio. The discrimination criteria for that linear model assumes low complexity scores below the line having a -45˚ slope and crossing data distribution 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.

Sequence assembly and taxonomic assignment

Clean and filtered MiSeq reads were assembled using as parameters 90% identity over a minimum of 50% of the read total length in CLC Genomics Workbench 4.4 (CLC bio USA, Cambridge, MA) [27]. Afterwards, contigs longer than 100 bp were queried for sequence similarity using BLASTN and BLASTX (NCBI-BLAST [28]) against the NCBI complete viral genomes database [29,30], the viral division of the GenBank nucleotide database [31,32], and viral proteins from UniProt [33]. The species nomenclature and classification followed NCBI Taxonomy database standards and the basic Baltimore classification. The alignments reported by BLAST (High-scoring Segment Pairs, HSPs) were required to have an E-value lower than 10^{-5} and a minimum length of 100 bp to be considered for taxonomical assessment. On the basis of the best BLAST results and a 90% coverage cut-off, the sequences were classified into their most likely taxonomic groups of origin.

Phylogenetic analysis

For *Anelloviridae*, phylogenetic trees were constructed based on the complete ORF1 region (with 75 reference sequences and a length alignment of 2551 bp), once contigs were properly aligned and trimmed. All the representative members of this family reported in humans were included as reference strains. Additionally, we also included some contigs longer than 1,500 bp that overlapped a large segment of ORF1 or a region upstream for individual trees. We compared each tree with the main tree generated from the reference strains to confirm equivalent distribution of species. In this manuscript, the following notation criteria were applied to name sequences on the phylogenetic trees: sequences covering ORF1, partially or not, were assigned to a number; contigs having some part outside ORF1 were identified with letters. For *Hepeviridae*, of the sequences mapped over the genome, we clipped the region that was present in all the sequence contigs under consideration. Then, the clipped region alignment was refined and some gaps were manually curated after visual inspection to improve the resulting alignment score. A reference phylogenetic tree was calculated from an alignment of 7483 bp with 22 known complete genomic sequences (19 of the genotype 3) as previously described [34]. Partial contig sequences aligning to a given particular region produced an equivalent tree. Those sequences were manually placed in the main tree according to the corresponding branches position on the equivalent trees, yet they are shown on the main reference tree as numbers or letters next to reference sequence identifier. All the alignments were produced by Geneious 10® as well as the phylogenetic trees, which were computed using the neighbour-joining method under the Jukes Cantor model. The robustness of the trees was assessed by bootstrap analysis of 1000 replicates each; finally, the branches are proportional to the corresponding phylogenetic distance.

Ethical statement

The study has been approved by the corresponding ethical committee: ethical committee on clinical investigation and research projects of the Hospital Universitari Vall D'Hebron

(N° 185; date: 4/2/2011). Serum samples were pooled at the hospital and for this study we do not have information on the identity of the patients.

Results

Nine libraries, consisting of 62 serum samples (32 of patients with unknown hepatitis, 10 of known HEV infections and 20 heathy volunteers), were obtained and sequenced using paired-end 300-base runs on the Illumina MiSeq platform, generating a total of 48 million reads (see Table 1 for a summary of the sequencing statistics for individual pools). Raw reads were binned by pool-based library barcodes and quality-filtered, leaving 30.5 million high-quality reads, which were assembled *de novo* within each pool subset. The resulting sequence contigs and singletons were compared to NCBI complete viral genomes, the viral division of the GenBank nucleotide database, and viral proteins retrieved from UniProt. Most of the viral sequences detected were related to the *Anelloviridae*, *Astroviridae*, *Caliciviridae*, *Hepeviridae*, *Flaviviridae* and *Retroviridae* families (Fig 1); those near-to-complete or partial genomes were characterized and are described in the following sections.

Volunteer samples that were analysed in the Healthy pools, in duplicate, show similar number of reads, and contigs. Additionally, the same families were found in those replicates, demonstrating that those results are highly consistent across samples (Table 1 and Fig 1).

Hepeviridae

A total of 27 contigs were matched to sequences of the *Hepeviridae* family. The HEV and Ai+ImSP pools produced sequences related to this family. A total of 76.1% (5,508 of 7,238 bp) of the HEV genome was sequenced from the HEV pool, with an average pairwise identity of 85.5% against the genotype 3 HEV (AF082843, Reference sequence genotype 3 ICTV). To identify the genotypes present in the pools and because metagenomics amplified different regions of the genome at random, individual phylogenetic trees were computed from contigs mapping over the same reference genome locations. The individual trees were compared to a reference species tree based on the reference-genomic sequences. Contigs that produced trees similar to the reference are marked in Fig 2 using numeric indexes, and information about each of those contigs is displayed on Table 2. On this table each contig is identified by its name (Contig ID), the contig length, its alignment identity percent to the homologous sequence from the blast HSPs, and confidence bootstrap value of the branch where it is placed on the corresponding phylogenetic tree. We were able to generate phylogenetic trees similar to the reference for eighteen contigs (the individual trees are available in S1 Supporting Information). Fifteen contigs from the HEV pool aligned to genotype 3f or closely related genotypes. The three contigs from the Ai+ImSP pool aligned with genotype 3a.

Anelloviridae

A total of 3,286 contigs matched sequences from the *Anelloviridae* family. All the pools produced sequences related to this family; however, the number of contigs was significantly higher in the pools with signs of hepatitis compared to the healthy pools (Wilcoxon rank-sum test, $p = 0.009$) and much more abundant in the Ai+ImSP pool (Fig 1). Contigs completely covering the ORF1 region of *Anelloviridae* family—or longer than 1,500 bp and overlapping this region—were found in the male A (less than 48 years old), female, HEV, and Ai+ImSP pools. Those particularly long sequences were used to build a phylogenetic tree to obtain a more accurate characterization of the species (Fig 3 and Table 3). The main members detected were Torque Teno Viruses (TTV—genus *Alphatorquevirus*) 1, 5, 10, 11, 13, 16, 18, 19, SEN virus H, Torque Teno Mini Viruses (TTMV—genus *Betatorquevirus*) 5, 9 and 18, Torque Teno Midi Viruses

Table 1. Summary of the sequences produced for each pool of serum samples in the sequencing experiment. All read counts correspond to total values, and the paired-reads real counts are half the values shown in the table. PE: paired-end reads; SE: single-end reads.

Pool ID	Number of samples	Raw Reads (PE MiSeq)	Clean Reads (PE + SE)	Contigs (after assembly)
Male A	8	5,255,854	3,614,220 + 6,928	43,188
Male B	8	2,669,124	1,769,992 + 4,738	19,000
Female	8	12,029,238	7,470,502 +18,074	83,518
Ai+ImSP	8	6,000,606	3,887,728 + 197,320	5,889
HEV	10	8,145,076	5,769,136 + 2,025	13,060
Healthy A.1	10	3,413,928	1,873,370 + 286	189,820
Healthy A.2		3,588,692	1,796,830 +298	166,167
Healthy B.1	10	3,457,150	2,119,224 +250	227,469
Healthy B.2		3,494,586	1,934,704 + 4	185,359

https://doi.org/10.1371/journal.pone.0185911.t001

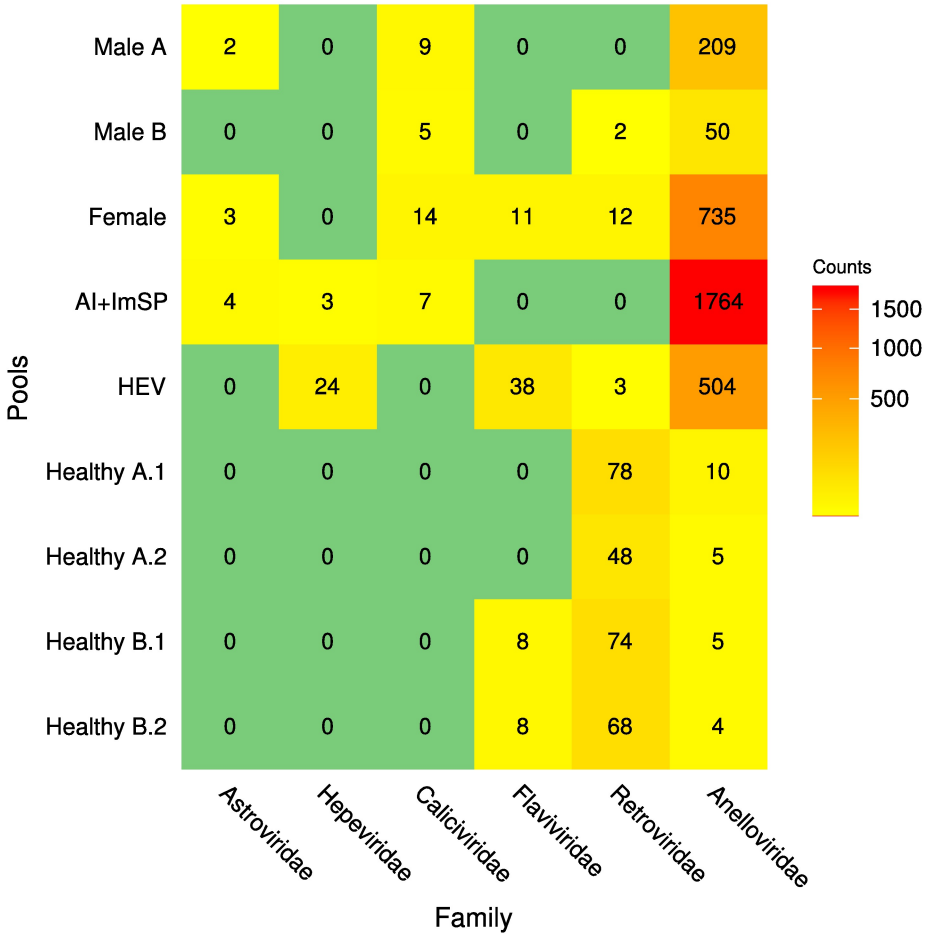


Fig 1. Heatmap describing number of contigs identified in each pool after their characterization and classification into taxonomic groups. Rows correspond to pooled samples whilst columns to families mapped at least to one sample. Numbers within each cell represent the number of sequences that had at least a positive BLAST hit to into known species and passed all the selection criteria. The colours range from yellow to red (low to high abundance respectively); green means that sequences were not detected for that group.

https://doi.org/10.1371/journal.pone.0185911.g001

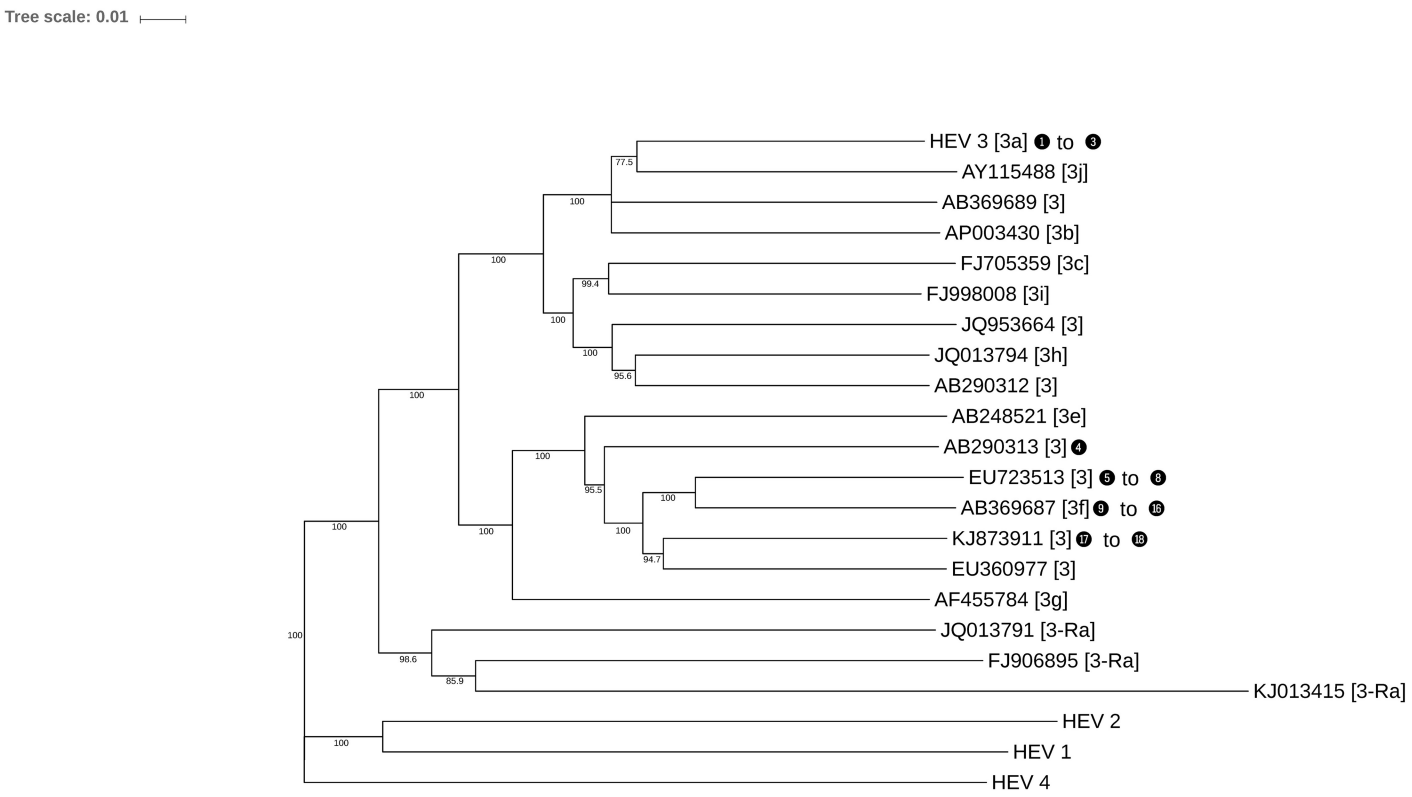


Fig 2. Phylogenetic tree of *Hepeviridae* based on complete genomes, including the main members of genotype 3. Numbers in blank bullets correspond to contigs identified in the HEV and Ai+ImSP pools (see Table 2); they are located beside the reference sequence where specific individual alignments of sequenced fragments over the same region in the reference sequences generated an equivalent tree topology (further results available from S1 Supporting Information). Labels within the square brackets define the species subtype. Small numbers on the tree branches show the bootstrap score of those branches.

https://doi.org/10.1371/journal.pone.0185911.g002

(TTMDV—genus *Gammatorquevirus*) 1, MDJN47, MDJN97, and other unclassified anelloviruses: TTV P19-3 (KT163917), TTV S72 (KP343839), TTV P1-3 (KT163877), TTV P13-4 (KT163899), TTMV Emory1 (KX810063), TTV S97 (KP343864), TTMV LY3 (JX134046), TTV S66 (KP343833), TTV S69 (KP343836), TTV S45 (KF545591), TTV P9-6 (KT163891), TTV S80 (KP343847), and TTV S57 (KP343824). Furthermore, contigs matching to the last two reference sequences do not belong to the three known genera of *Anelloviridae* previously identified in humans; thus, it seems they define a new cluster/genus for this family. Moreover, 60% (19/32) of the longest contigs have less than 80% identity to the already described sequences from the NCBI database. Table 3 shows the contigs that were considered for this phylogenetic analysis; each contig is identified by its name (contig ID), sequence length in bp, alignment identity percent to the homologus sequence from the BLAST HSPs, and confidence bootstrap value of the branch where it is placed on the corresponding phylogenetic tree (individual trees are provided in the S2 Supporting Information). Fewer and shorter contigs were found in the pools from healthy individuals in comparison with the other pools (median of 300 bp); they correspond to TTV 1, 19 and TTMV 6.

Caliciviridae

A total of 35 contigs between 200 and 654 bp aligned to the *Caliciviridae* family. They were found in the male A and B, female and Ai+ImSP pools. No sequences of this family were

Table 2. Summary of similarity searches for those detected from the HEV and Ai+ImSP pools. The first column corresponds to the numbers in the black bullets shown on some of the branches of the *Hepeviridae* phylogenetic tree from Fig 2.

Code	Pool	Contig ID	Length	%Identity	Bootstrap
1	AI+IMSP	contig_953	992	91.63	100.0
2	AI+IMSP	contig_1893	686	91.40	97.4
3	AI+IMSP	contig_3606	412	91.02	84.1
4	HEV	contig 3810	573	86.83	60.0
5	HEV	contig_2453	1,416	91.05	68.4
6	HEV	contig_533	590	90.51	70.0
7	HEV	contig_1542	575	89.04	76.7
8	HEV	contig 749	541	88.54	51.3
9	HEV	contig_6571	1,572	91.35	99.9
10	HEV	contig_747	1,415	88.30	74.6
11	HEV	contig _3424	1,032	90.31	83.3
12	HEV	contig_6979	944	91.58	100.0
13	HEV	contig_7146	929	89.26	95.0
14	HEV	contig_8370	557	93.33	96.0
15	HEV	contig 10460	314	86.29	95.4
16	HEV	contig 3914	297	87.21	65.2
17	HEV	contig_1444	1,534	87.11	98.3
18	HEV	contig 3007	333	89.33	86.7

https://doi.org/10.1371/journal.pone.0185911.t002

detected in healthy volunteer pools. All contigs were assigned to sapovirus Hu/Nagoya/NGY (AB775659), genogroup 5 strain 2 (GV.2), with identities varying between 97% and 100%. Those contigs map over several regions of the non-structural protein and major structural protein, including eleven that aligned to a partial capsid fragment.

Astroviridae

As few as eight contigs between 214 and 493 bp long matched the *Astroviridae* family. They were found in the male A (less than 48 years old), female, and Ai+ImSP pools. No sequences of this family were detected in the healthy-volunteers pools. These contigs correspond to a recently discovered astrovirus, clade VA strain 3 (VA3, also known as HMO-C) (7 matching JX857868, 1 matching JX083288), with identities ranging from 97% to 100%.

Flaviviridae

A total of 65 contigs between 219 and 2778 bp matched the *Flaviviridae* family. They were found in the female, Ai+ImSP, and healthy B1 and B2 pools. All the sequences aligned to several entries of GB virus C from GenBank, with identities between 97% and 100%.

Retroviridae

In this case, 285 contigs between 300 and 1,032 bp were assigned to the *Retroviridae* family. They were found in the male B (more than 48 years old), female, and HEV pools and in all healthy pools. All the sequences matched several entries of human endogenous retrovirus type K and HCML-ARV with identities greater than 70%.

The raw sequencing data used to perform this analysis along with the FASTQ file are located in the NCBI Sequence Read Archive; BioProject (PRJNA379441).

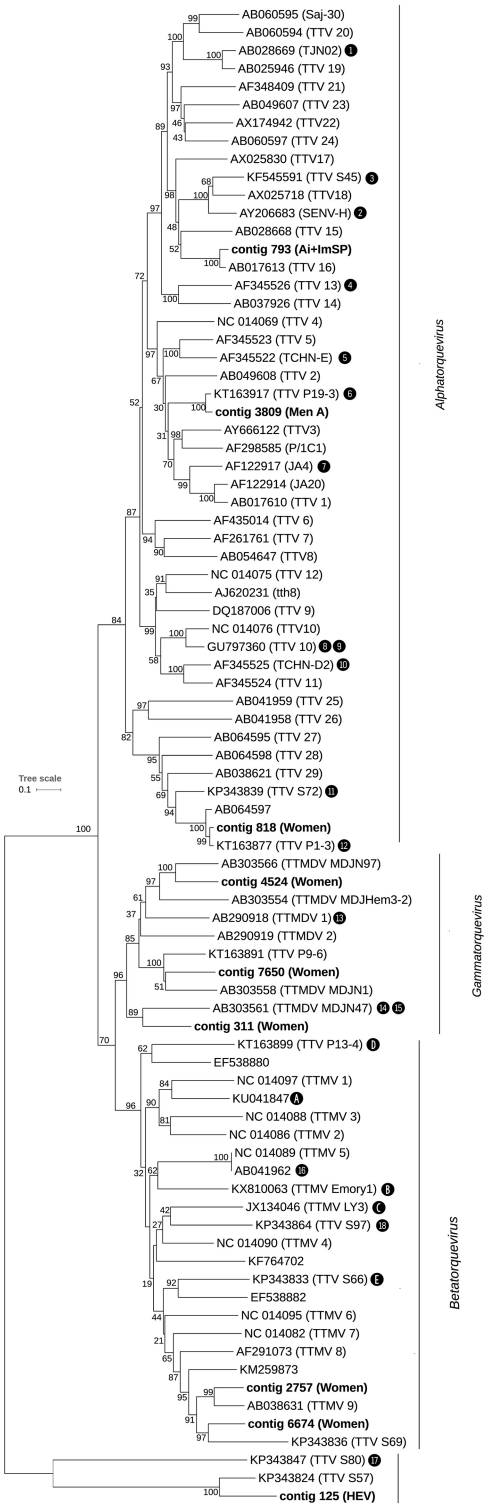


Fig 3. Phylogenetic tree for the *Anelloviridae* family based on ORF1 region and including only contigs that fully overlap with that region. Numbers and letters within black bullets refer to contigs longer than 1,500 bp (see Table 3) that partially aligned with ORF1 or with the ORF1 upstream region, respectively. See Fig 2 for further details about notation used in this tree.

https://doi.org/10.1371/journal.pone.0185911.g003

Table 3. Summary information for contigs longer than 1,500 bp that were found in the pooled samples and assigned to the *Anelloviridae* family. The number and letter codes from the first column (Code) correspond to those in the blank bullets shown on some of the branches of the phylogenetic tree from Fig 3. Those without codes were placed directly on the tree, as they defined new branches.

Code	Sample	Contig ID	Length (bp)	Sequence Name	% Identity	Bootstrap
	Male A	contig_3809	2,798	TTV P19-3	92.8	100.0
D	Male A	contig_1199	1,512	TTV P13-4	69.3	68.0
10	Male A	contig_7929	1,875	TCHN-D2/TTV 11	79.2	100.0
12	Female	contig_1	1,548	TTV P1-3	92.4	100.0
11	Female	contig_129	1,506	TTV S72	79.8	100.0
14	Female	contig_16376	1,513	TTMDV MDJN47	68.3	91.9
	Female	contig_2757	2,142	TTMV 9	76.6	99.0
	Female	contig_4524	1,977	TTMDV MDJN97	72.1	100.0
1	Female	contig_5911	1,500	TJN2/TTV19	88.9	100.0
A	Female	contig_626	1,503	TTMV 18	89.7	100.0
	Female	contig_6674	2,381	TTV S69	71.2	97.0
	Female	contig_818	2,264	TTV P1-3	95.1	99.0
7	Female	contig_9035	2,243	JA4	94.0	100.0
4	Female	contig_1475	1,899	TCHN-A	87.5	100.0
13	Female	contig_1946	1,644	TTMDV1	71.6	100.0
2	Female	contig_268	1,951	SENV-H	90.1	100.0
	Female	contig_311	2,086	TTMDV MDJN47	66.6	89.0
8	Female	contig_6533	1,530	TTV10	83.3	100.0
	Female	contig_7650	1,730	TTV P9-6	69.6	51.0
17	AI+IMSP	contig_1199	1,586	TTV S80	66.8	99.8
	AI+IMSP	contig_793	2,367	TTV 16	90.8	100.0
15	AI+IMSP	contig_1013	1,687	TTMDV MDJN47	66.0	95.8
9	AI+IMSP	contig_1709	1,832	TTV 10	85.7	100.0
18	AI+IMSP	contig_2151	1,985	TTV S97	65.6	79.0
B	HEV	contig_118	1,781	TTMV Emory1	68.6	68.0
3	HEV	contig_2837	1,845	TTV S45	86.6	100.0
	HEV	contig_125	2,303	TTV S57	70.1	100.0
C	HEV	contig_2	1,923	TTMV LY3	66.6	79.0
5	HEV	contig_236	1,643	TTV TCHN-E	72.0	99.9
6	HEV	contig_2366	1,514	TTV P19-3	93.8	100.0
16	HEV	contig_506	2,046	TTMV 5	73.7	96.0
E	HEV	contig_66	1,904	TTV S66	71.9	49.0

https://doi.org/10.1371/journal.pone.0185911.t003

Discussion

The aim of this study was to investigate viruses infecting patients diagnosed with acute hepatitis. Different groups of patients presenting with acute hepatitis but without serological infection markers of the most common viral hepatitis were studied to determine possible causal agents of non-A-E hepatitis. Our findings demonstrate the presence of a high variety of viral sequences in pools of patients with hepatitis of unknown aetiology.

HEV viruses were detected in two pools (HEV and Ai+ImSP). We found a variety of contigs related to genotype 3f in HEV pools. Genotype 3f has been described in hepatitis outbreaks in Catalonia [35], Spain [36] and the south of France [37]. This strain has also been related to swine and wild boar consumption, which can be considered a food-borne and an emerging zoonotic infection [35,38,39]. Individual samples from the Ai+ImSP pool were re-analysed afterwards by nPCR, and one patient was identified as HEV-positive in this second round,

which would explain the presence of HEV contigs in this pool. Metagenomics approaches have the advantage of identifying more than one genotype in the pools; this facilitates description of traces of possible multiple infections in a single sample.

We have found at least three different kinds of *Anelloviridae* contigs: a) contigs that match previously characterized sequences; b) contigs that are closely related to unclassified sequences; and, c) contigs poorly related to classified and unclassified sequences (potential new viruses). The demarcation criteria of the genus establish a cut-off value of 35% nucleic-acid identity in the ORF1 region. Due to the number of quasispecies discovered in this family [40], it is difficult to establish a clear cut-off at the species level.

We also describe in this paper viruses that have been previously associated with hepatitis such as TTV-1, 11, 16 and SEN virus H [14,41]; other viruses have been recently described in serum samples from HIV patients (P19-3, P13-4, P9-6, P1-3); yet other sets were described in patients with various conditions, including lymphocytic leukaemia (TTV 10) [42], gingival periodontitis (TTMV 18) [43], haemophilia (TTMDV MDJN47 and MDJN97) [44] and in pregnant women whose offspring developed leukaemia and lymphomas (TTV S45, S57, S66, S69, S72, S80 and S97) [45].

Metagenomics analyses are driving the discovery of new potential sequences in this family; Bzhalava et al. (2016) described for first time a group of sequences detected from human samples, spawning a new branch of the *Anelloviridae* family. We found two contigs (125 and 1199) falling into this new potential genus of *Anelloviridae*, yet they have less than 70% of identity to those sequences, which were described in serum samples from pregnant women. Such results suggest that there will be more viruses within this family that have not yet been identified.

TTV-1, the first member identified in the *Anelloviridae* family, was reported in hepatitis patients in whom no causative agents were detected [12]. This family includes three genera that have been identified in humans: *Alphatorquevirus* (TTV), *Betatorquevirus* (TTMV), and *Gammatorquevirus* (TTMDV) [46]. However, the role of those viruses in hepatitis or in other diseases remains uncertain [14,40,47]. Numerous recent studies have demonstrated a prevalence between 5 and 90% in the blood of the general population, depending on the geographic region [40]. Moreover, the genetic diversity among anelloviruses is far greater than it is within any other group of ssDNA viruses. The considerable genetic heterogeneity is exemplified by the large number of highly divergent sequences being identified in this family. There are at least 41 species infecting humans that are recognized by the ICTV based on the ORF1 region [46]. Some viruses, such as TTV 1, 12, 13, 16, SEN virus D and H, have been considered potential causal agents of hepatitis [14,48–50].

Unfortunately, anelloviruses cannot be propagated *in vitro* due to the lack of compatible cell systems. However, they have a high *in vivo* replication capacity. Infection with TTV is characterized by persistent lifelong viremia in humans, with circulation levels of up to 10^6 genomic copies/ml in the general population [14,40]. TTV replicates in the liver and is excreted at high levels in bile and faeces [51]. Additionally, other studies have shown that this virus does not have a particular tropism [40,52]. Metagenomic analyses have also shown that TTV is a common finding in several sample types [53]. For that reason, determining the causative factors of illness can be difficult.

An increased number of contigs aligning to anelloviruses was observed in this study, however, these findings not necessarily may support the hypothesis that these viruses are the causative agents. Previous studies have suggested titres of TTV in plasma as an indicator of immune status [54]. Another study showed that anellovirus load in plasma increases substantially during immunosuppressive therapy and in immunocompromised patients [55]. Shotgun sequencing from plasma samples that were collected over several months post-transplantation also

revealed that viral loads increased, whereas the bacterial composition remained unchanged [56].

The results described in this study also show the presence of sapovirus strain GV.2 in all the pools of patients with clinical hepatitis of unknown aetiology. This strain has been recently characterized from faecal samples from a suspected foodborne gastroenteritis outbreak in Japan using a metagenomics sequencing approach [57]. Partial fragments of that virus were described early from another gastroenteritis outbreak in Italy [58], in river water from Barcelona (the same region where this study was conducted) [59], and in wastewater from Japan [60], suggesting prevalent circulation of this virus around the world. Sapovirus are positive-sense single-stranded RNA viruses from the family *Caliciviridae*. Members of this family are known to cause gastroenteritis with self-limited infections and low mortality rates; severe infections or serious clinical complications are usually reported in immunocompromised patients [61]. Further research would be required to analyse the possible pathogenic role of sapovirus GV.2 in our study.

Few contigs of the *Astroviridae* family were detected in this work. Astrovirus VA3 was identified in most of the pools of hepatitis of unknown aetiology. However, those contigs were less abundant and shorter than the sapovirus contigs. The first description of astrovirus VA3 was from the stool of paediatric patients with diarrhoea from India [62], and it was later completely sequenced [63]. This virus has also been described in stools from southern China [64], Kenya, and the Gambia [65]. However, the role of this virus in health and disease remain largely unknown.

The potential pathogenic role of sapovirus GV.2 and astrovirus VA3 in blood remains still uncertain. Although astroviruses and sapoviruses are considered gastrointestinal pathogens, viral RNA and infectious particles have been recovered from extraintestinal organs in both animals and humans. Examples in animals implicate astroviruses as the cause of hepatitis in ducks [66] and the isolation of murine astroviruses in mouse liver [67]. With respect to sapovirus less information is available; an isolation of sapovirus in a liver of a spotted hyena [68]. Our results suggest that the presence of these viruses in pools from patients with non A-to-E hepatitis, including the AI+ImSp pool, merits further research, since there is no previous evidence relating those viruses to hepatitis.

GB virus C, also known as pegivirus or hepatitis G virus, is a human virus of the *Flaviviridae* family that is structurally and epidemiologically closest to hepatitis C virus [13]. Most GBV-C infections appear to be asymptomatic, transient, and self-limiting, with slight or no elevation of ALT levels. Those infections are rarely identified and very difficult to evaluate. The role of GBV-C in the aetiology of hepatitis has not been fully established [69]. Moreover, it is commonly reported in metagenomics studies [53], suggesting its limited role in the development of illness, including hepatitis. We have detected this virus in one healthy pool and in a hepatitis pool; our results support the hypothesis that this species may be widely distributed within the population.

Human endogenous retroviruses (HERVs) are remnants of germ-line retrovirus integration and are considered functionally defective [70]. They have been described in metagenomics studies at high levels [55,70] without association with any particular pathology [71]. Our findings support previous results pointing out that this virus is present in healthy people.

It is important to recognize that the use of serum samples to describe the virome may have some minor limitations as a decreased sensitivity to detect integrated proviruses (e.g. HIV-1), episomal viruses (e.g. herpesviruses) [72]. Furthermore, giant viruses may also be under-represented due to the filtration process [73]. However, serum samples predominantly contain host DNA which can also affect the sensitivity of viral detection [74]; if host and viral NA cannot be easily separated, the resulting fraction of viral sequences relative to the host DNA would be

extremely low [53]. Pretreatments protocols for viral enrichment have to be taken into consideration in order to get a better approximation to the whole virome and the interaction between virus population in future studies.

Conclusions

In summary, metagenomics was applied in this study to detect a broad spectrum of viral species based on sequences found in pooled samples, including HEV in pools of patients with confirmed HEV; these samples allowed the characterization of the most prevalent genotypes. Additionally, we were able to identify a diverse population of anelloviruses, including novel undescribed sequences, in patients with acute hepatitis of unknown aetiology. Furthermore, sapovirus GV.2 and astrovirus VA3, viruses recently reported as cause of gastroenteritis, were also found exclusively in those pools. We did not attempt to determine causality or to describe epidemiologic results; our purpose was to characterize the virome of patients diagnosed with hepatitis to describe new potential causal agents. The role of these viruses as possible causal agents of hepatitis of unknown aetiology remains open to further studies. Finally, reproducibility between replicates in the pools of healthy volunteers supports the consideration of the metagenomics as a robust detection method for viral species. Metagenomics analyses offer unprecedented possibilities for diagnostics, characterization and identification of possible co-infections of rare and novel viruses that will be relevant to understanding the aetiology of current pathologies without known causative agents.

Supporting information

S1 Supporting Information. Individual phylogenetic trees computed from contigs over reference genome locations in HEV.
(DOCX)

S2 Supporting Information. Individual phylogenetic trees computed from contigs over reference genome locations in *Anelloviridae* family.
(DOCX)

Acknowledgments

The study reported here was partially funded by the Programa RecerCaixa 2012 (ACUP-00300), AGL2011-30461-C02-01/ALI and AGL2014-55081-R from Spanish MINECO. 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, Eloy Gonzales-Gustavson is a fellow of the Peruvian Government; Natalia Timoneda is a fellow of the Spanish Ministry of Science and Xavier Fernandez-Cassi was a fellow of the Catalan Government "AGAUR" (FI-DGR).

Author Contributions

Conceptualization: Eloy Gonzales-Gustavson, R. Girones.

Formal analysis: Eloy Gonzales-Gustavson, N. Timoneda, J. F. Abril, R. Girones.

Funding acquisition: R. Girones.

Investigation: Eloy Gonzales-Gustavson, N. Timoneda, X. Fernandez-Cassi, A. Caballero, J. F. Abril, M. Buti, R. Girones.

Methodology: Eloy Gonzales-Gustavson, X. Fernandez-Cassi, A. Caballero, F. Rodriguez-Frias, R. Girones.

Project administration: R. Girones.

Resources: Eloy Gonzales-Gustavson, M. Buti, F. Rodriguez-Frias.

Software: Eloy Gonzales-Gustavson, N. Timoneda.

Supervision: J. F. Abril, R. Girones.

Visualization: Eloy Gonzales-Gustavson.

Writing – original draft: Eloy Gonzales-Gustavson.

Writing – review & editing: Eloy Gonzales-Gustavson, N. Timoneda, J. F. Abril, R. Girones.

References

1. Previsani N, Lavanchy D, Siegl G. Hepatitis A. In: Mushahwar Isa K, editor. Viral Hepatitis: Molecular Biology, Diagnosis, Epidemiology and Control. 1st ed. Amsterdam: ElSevier; 2004. pp. 1–30.
2. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012; 380: 2095–2128. https://doi.org/10.1016/S0140-6736(12)61728-0 PMID: 23245604
3. Manka P, Verheyen J, Gerken G, Canbay A. Liver Failure due to Acute Viral Hepatitis (A-E). Visc Med. 2016; 80–85. https://doi.org/10.1159/000444915 PMID: 27413724
4. Alter MJ, Gallagher M, Morris TT, Moyer LA, Meeks EL, Krawczynski K, et al. Acute Non-A–E Hepatitis in the United States and the Role of Hepatitis G Virus Infection. N Engl J Med. Massachusetts Medical Society; 1997; 336: 741–746. https://doi.org/10.1056/NEJM199703133361101 PMID: 9052651
5. Cacopardo B, Nunnari G, Berger A, Doer HW, Russo R. Acute non A-E hepatitis in eastern Sicily: the natural history and the role of hepatitis G virus. Eur Rev Med Pharmacol Sci. 2000; 4: 117–121. Available: http://www.ncbi.nlm.nih.gov/pubmed/11710508 PMID: 11710508
6. Chu C-M, Lin S-M, Hsieh S-Y, Yeh C-T, Lin D-Y, Sheen I-S, et al. Etiology of sporadic acute viral hepatitis in Taiwan: The role of hepatitis C virus, hepatitis E virus and GB virus-C/hepatitis G virus in an endemic area of hepatitis A and B. J Med Virol. John Wiley & Sons, Inc.; 1999; 58: 154–159. https://doi.org/10.1002/(SICI)1096-9071(199906)58:2<154::AID-JMV9>3.0.CO;2-E
7. Delic D, Mitrovi N, Radovanovi A. Epidemiological characteristics and clinical manifestations of acute non-A-E hepatitis. Vojnosanit Pregl. 2010; 67: 903–909. PMID: 21268515
8. Tassopoulos NC, Papatheodoridis G V, Delladetsima I, Hatzakis A. Clinicopathological features and natural history of acute sporadic non-(A-E) hepatitis. J Gastroenterol Hepatol. 2008; 23: 1208–1215. https://doi.org/10.1111/j.1440-1746.2008.05454.x PMID: 18554239
9. Yeh C-T, Tsao M-L, Lin Y-C, Tseng I-C. Identification of a novel single-stranded DNA fragment associated with human hepatitis. J Infect Dis. 2006; 193: 1089–97. https://doi.org/10.1086/501474 PMID: 16544249
10. Van Thiel DH, Gavaler JS, Baddour N, Friedlander L, Wright HI. Treatment of putative non-A, non-B, non-C hepatitis with alpha interferon: a preliminary trial. J Okla State Med Assoc. UNITED STATES; 1994; 87: 364–368.
11. Simons JN, Leary TP, Dawson GJ, Pilot-Matias TJ, Muerhoff a S, Schlauder GG, et al. Isolation of novel virus-like sequences associated with human hepatitis. Nat Med. 1995; 1: 564–569. https://doi.org/10.1038/nm0695-564 PMID: 7585124
12. Nishizawa T, Okamoto H, Konishi K, Yoshizawa H, Miyakawa Y, Mayumi M. A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology. Biochem Biophys Res Commun. 1997; 241: 92–7. https://doi.org/10.1006/bbrc.1997.7765 PMID: 9405239
13. Chivero ET, Stapleton JT. Tropism of human pegivirus (Formerly known as GB virus C/hepatitis G virus) and host immunomodulation: Insights into a highly successful viral infection. J Gen Virol. 2015; 96: 1521–1532. https://doi.org/10.1099/vir.0.000086 PMID: 25667328
14. Okamoto H. History of discoveries and pathogenicity of TT viruses. In: de Villiers E-M, zur Hausen H, editors. TT viruses—The still elusive human pathogens. First. Berlin: Springer; 2009. pp. 1–20. https://doi.org/10.1007/978-3-540-70972-5_1

15. Conrad M, Knodell R. Viral hepatitis—1975. JAMA. 2014; 312: 654. Available: http://dx.doi.org/10.1001/jama.2013.279664 PMID: 25117146
16. Kaufman B, Gandhi SA, Louie E, Rizzi R, Illei P. Herpes simplex virus hepatitis: case report and review. Clin Infect Dis. 1997; 24: 334–338. PMID: 9114181
17. Bihari C, Rastogi a, Saxena P, Rangegowda D, Chowdhury a, Gupta N, et al. Parvovirus B19 Associated Hepatitis. Hepat Res Treat. 2013; 2013: 472027. https://doi.org/10.1155/2013/472027 PMID: 24232179
18. Kawashima N, Muramatsu H, Okuno Y, Torii Y, Kawada J, Narita A, et al. Fulminant adenovirus hepatitis after hematopoietic stem cell transplant: Retrospective real-time PCR analysis for adenovirus DNA in two cases. J Infect Chemother. Elsevier Ltd; 2015; 21: 857–863. https://doi.org/10.1016/j.jiac.2015.08.018 PMID: 26423689
19. Mateos ME, López-Laso E, Pérez-Navero JL, Peña MJ, Velasco MJ. Successful response to cidofovir of adenovirus hepatitis during chemotherapy in a child with hepatoblastoma. J Pediatr Hematol Oncol. 2012; 34: e298–300. https://doi.org/10.1097/MPH.0b013e318266ba72 PMID: 22935664
20. Chu C-M, Lin D-Y, Yeh C-T, Sheen I-S, Liaw Y-F. Epidemiological Characteristics, Risk Factors, and Clinical Manifestations of Acute Non-A±E Hepatitis. J Med Virol J Med Virol. 2001;65.
21. Ogilvie L, Jones B. The human gut virome: a multifaceted majority. Front Microbiol. 2015; 6: 918. https://doi.org/10.3389/fmicb.2015.00918 PMID: 26441861
22. Hugenholtz P, Tyson GW. Metagenomics. Nature. 2008; 455: 481–483. https://doi.org/10.1038/455481a PMID: 18818648
23. 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. https://doi.org/10.1073/pnas.242579699 PMID: 12429852
24. Wang D, Urisman A, Liu YT, Springer M, Ksiazek TG, Erdman DD, et al. Viral discovery and sequence recovery using DNA microarrays. PLoS Biol. 2003; 1. https://doi.org/10.1371/journal.pbio.0000002 PMID: 14624234
25. Hannon G. FASTX-Toolkit [Internet]. 2015. http://hannonlab.cshl.edu/fastx_toolkit/index.html
26. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 PMID: 24695404
27. CLC Bio. CLC Genomics Workbench [Internet]. 2008. https://www.qiagenbioinformatics.com/products/clc-genomics-workbench/
28. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215: 403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2 PMID: 2231712
29. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. 1997; 25: 3389–3402. https://doi.org/10.1093/nar/25.17.3389 PMID: 9254694
30. Brister JR, Ako-Adjei D, Bao Y, Blinkova O. NCBI viral Genomes resource. Nucleic Acids Res. 2015; 43: D571–D577. https://doi.org/10.1093/nar/gku1207 PMID: 25428358
31. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res. 2016; 44: D67–D72. https://doi.org/10.1093/nar/gkv1276 PMID: 26590407
32. NCBI. National Center for Biotechnology Information (NCBI) assembled genomes [Internet]. 2016. ftp://ftp.ncbi.nlm.nih.gov/genomes/
33. Bateman A, Martin MJ, O'Donovan C, Magrane M, Apweiler R, Alpi E, et al. UniProt: A hub for protein information. Nucleic Acids Res. 2015; 43: D204–D212. https://doi.org/10.1093/nar/gku989 PMID: 25348405
34. Smith DB, Simmonds P, Izopet J, Oliveira-Filho EF, Ulrich RG, Johne R, et al. Proposed reference sequences for hepatitis E Virus subtypes. J Gen Virol. 2016; 97: 537–542. https://doi.org/10.1099/jgv.0.000393 PMID: 26743685
35. Riveiro-Barciela M, Mínguez B, Gironés R, Rodriguez-Frías F, Quer J, Buti M. Phylogenetic demonstration of hepatitis E infection transmitted by pork meat ingestion. J Clin Gastroenterol. 2015; 49: 165–8. https://doi.org/10.1097/MCG.0000000000000113 PMID: 24637729
36. Rivero-Juarez A, Frias M, Martinez-Peinado A, Risalde MA, Rodriguez-Cano D, Camacho A, et al. Familial Hepatitis E Outbreak Linked to Wild Boar Meat Consumption. Zoonoses Public Health. 2017; 1–5.
37. Legrand-Abravanel F, Mansuy JM, Dubois M, Kamar N, Peron JM, Rostaing L, et al. Hepatitis E virus genotype 3 diversity, France. Emerg Infect Dis. 2009; 15: 110–114. https://doi.org/10.3201/eid1501.080296 PMID: 19116067

38. Banks M, Bendall R, Grierson S, Heath G, Mitchell J, Dalton H. Human and porcine hepatitis E virus strains, United Kingdom. Emerg Infect Dis. 2004; 10: 953–955. https://doi.org/10.3201/eid1005.030908 PMID: 15200841

39. Meng XJ, Purcell RH, Halbur PG, Lehman JR, Webb DM, Tsareva TS, et al. A novel virus in swine is closely related to the human hepatitis E virus. Proc Natl Acad Sci U S A. 1997; 94: 9860–9865. PMID: 9275216

40. Spandole S, Cimponeriu D, Berca LM, Mihaescu G. Human anelloviruses: an update of molecular, epidemiological and clinical aspects. Arch Virol. Austria; 2015; 160: 893–908. https://doi.org/10.1007/s00705-015-2363-9 PMID: 25680568

41. Luo K, He H, Liu Z, Liu D, Xiao H, Jiang X, et al. Novel variants related to TT virus distributed widely in China. J Med Virol. 2002; 67: 118–126. https://doi.org/10.1002/jmv.2200 PMID: 11920826

42. Chu C, Zhang L, Dhayalan A, Agagnina BM, Magli AR, Fraher G, et al. Torque Teno Virus 10 Isolated by Genome Amplification Techniques from a Patient with Concomitant Chronic Lymphocytic Leukemia and Polycythemia Vera. Mol Med. 2011; 17: 1. https://doi.org/10.2119/molmed.2010.00110 PMID: 21953418

43. Zhang Y, Li F, Shan T-L, Deng X, Delwart E, Feng X-P. A novel species of torque teno mini virus (TTMV) in gingival tissue from chronic periodontitis patients. Sci Rep. Nature Publishing Group; 2016; 6: 26739. https://doi.org/10.1038/srep26739 PMID: 27221159

44. Ninomiya M, Takahashi M, Shimosegawa T, Okamoto H. Analysis of the entire genomes of fifteen torque teno midi virus variants classifiable into a third group of genus Anellovirus. Arch Virol. 2007; 152: 1961–1975. https://doi.org/10.1007/s00705-007-1046-6 PMID: 17712598

45. Bzhalava D, Hultin E, Muhr LSA, Ekstrom J, Lehtinen M, de Villiers EM, et al. Viremia during pregnancy and risk of childhood leukemia and lymphomas in the offspring: nested case-control study. Int J Cancer. 2016; 138: 2212–2220. https://doi.org/10.1002/ijc.29666 PMID: 26132655

46. Biagini P, Bendinelli M, Hino S, Kakkola L, Mankertz A, Niel C, et al. Anelloviridae. In: King A, Adams M, Carstens E, Lefkowitz E, editors. Virus taxonomy, clasification and nomenclature of viruses. Ninth Repo. Davis: ElSevier Academic Press; 2012. pp. 331–341.

47. Hsiao KL, Wang LY, Lin CL, Liu HF. New phylogenetic groups of torque teno virus identified in eastern Taiwan indigenes. PLoS One. 2016; 11: 1–10. https://doi.org/10.1371/journal.pone.0149901 PMID: 26901643

48. Bostan N. Current and Future Prospects of Torque Teno Virus. J Vaccines Vaccin. 2013; 1–9. https://doi.org/10.4172/2157-7560.S1-004

49. Kakkola L, Bondén H, Hedman L, Kivi N, Moisala S, Julin J, et al. Expression of all six human Torque teno virus (TTV) proteins in bacteria and in insect cells, and analysis of their IgG responses. Virology. Elsevier Inc.; 2008; 382: 182–189. https://doi.org/10.1016/j.virol.2008.09.012 PMID: 18947848

50. Mi Z, Yuan X, Pei G, Wang W, An X, Zhang Z, et al. High-throughput sequencing exclusively identified a novel Torque teno virus genotype in serum of a patient with fatal fever. Virol Sin. 2014; 29: 112–118. https://doi.org/10.1007/s12250-014-3424-z PMID: 24752764

51. Ohbayashi H, Tanaka Y, Ohoka S, Chinzei R, Kakinuma S, Goto M, et al. TT virus is shown in the liver by in situ hybridization with a PCR-generated probe from the serum TTV-DNA. J Gastroenterol Hepatol. 2001; 16: 424–428. https://doi.org/10.1046/j.1440-1746.2001.02460.x PMID: 11354281

52. Okamoto H, Nishizawa T, Takahashi M. Torque teno virus (TTV): molecular virology and clinical implications. In: Mushawar IK, editor. Viral Hepatitis: Molecular Biology, Diagnosis, Epidemiology and Control. Amsterdan: Elsevier; 2004. pp. 241–251.

53. Delwart E. Viral metagenomics. Rev Med Virol. John Wiley & Sons, Ltd.; 2007; 17: 115–131. https://doi.org/10.1002/rmv.532 PMID: 17295196

54. Touinssi M, Gallian P, Biagini P, Attoui H, Vialettes B, Berland Y, et al. TT virus infection: prevalence of elevated viraemia and arguments for the immune control of viral load. J Clin Virol. 2001; 21: 135–41. Available: http://www.ncbi.nlm.nih.gov/pubmed/11378494 PMID: 11378494

55. Li L, Deng X, Linsuwanon P, Bangsberg D, Bwana MB, Hunt P, et al. AIDS alters the commensal plasma virome. J Virol. 2013; 87: 10912–5. https://doi.org/10.1128/JVI.01839-13 PMID: 23903845

56. Hofer U. Microbiome: anelloviridae go viral. Nat Rev Microbiol. Nature Publishing Group; 2014; 12: 4–5. https://doi.org/10.1038/nrmicro3192 PMID: 24336177

57. Shibata S, Sekizuka T, Kodaira A, Kuroda M, Haga K, Doan YH, et al. Complete Genome Sequence of a Novel GV.2 Sapovirus Strain, NGY-1, Detected from a Suspected Foodborne Gastroenteritis. Genome Announc. 2015; 3: 2–3.

58. Medici MC, Tummolo F, Albonetti V, Abelli LA, Chezzi C, Calderaro A. Molecular detection and epidemiology of astrovirus, bocavirus, and sapovirus in Italian children admitted to hospital with acute

gastroenteritis, 2008–2009. J Med Virol. United States; 2012; 84: 643–650. https://doi.org/10.1002/jmv.23231 PMID: 22337304

59. Sano D, Pérez-Sautu U, Guix S, Pintó RM, Miura T, Okabe S, et al. Quantification and genotyping of human sapoviruses in the Llobregat river catchment, Spain. Appl Environ Microbiol. 2011; 77: 1111–1114. https://doi.org/10.1128/AEM.01721-10 PMID: 21148702

60. Hansman GS, Sano D, Ueki Y, Imai T, Oka T, Katayama K, et al. Sapovirus in water, Japan. Emerg Infect Dis. 2007; 13: 133–135. https://doi.org/10.3201/eid1301.061047 PMID: 17370528

61. Oka T, Wang Q, Katayama K, Saif LJ. Comprehensive review of human sapoviruses. ClinMicrobiolRev. 2015; 28: 32–53. https://doi.org/10.1128/CMR.00011-14 PMID: 25567221

62. Finkbeiner S, Holtz L, Jiang Y, Rajendran P, Franz C, Zhao G, et al. Human stool contains a previously unrecognized diversity of novel astroviruses. Virol J. 2009; 6: 161. https://doi.org/10.1186/1743-422X-6-161 PMID: 19814825

63. Jiang H, Holtz LR, Bauer I, Franz CJ, Zhao G, Bodhidatta L, et al. Comparison of novel MLB-clade, VA-clade and classic human astroviruses highlights constrained evolution of the classic human astrovirus nonstructural genes. Virology. 2013; 436: 8–14. https://doi.org/10.1016/j.virol.2012.09.040 PMID: 23084422

64. Xiao J, Li J, Hu G, Chen Z, Wu Y, Chen Y, et al. Isolation and phylogenetic characterization of bat astroviruses in southern China. Arch Virol. 2011; 156: 1415–1423. https://doi.org/10.1007/s00705-011-1011-2 PMID: 21573690

65. Meyer CT, Bauer IK, Antonio M, Adeyemi M, Saha D, Oundo JO, et al. Prevalence of classic, MLB-clade and VA-clade Astroviruses in Kenya and The Gambia. Virol J. Virology Journal; 2015; 12: 78. https://doi.org/10.1186/s12985-015-0299-z PMID: 25975198

66. Liu N, Wang F, Shi J, Zheng L, Wang X, Zhang D. Molecular characterization of a duck hepatitis virus 3-like astrovirus. Vet Microbiol. Elsevier B.V.; 2014; 170: 39–47. https://doi.org/10.1016/j.vetmic.2014.01.026 PMID: 24560589

67. Yokoyama CC, Loh J, Zhao G, Stappenbeck TS, Wang D, Huang H V, et al. Adaptive Immunity Restricts Replication of Novel Murine Astroviruses. J Virol. 2012; 86: 12262–12270. https://doi.org/10.1128/JVI.02018-12 PMID: 22951832

68. Olarte-Castillo XA, Hofer H, Goller K V., Martella V, Moehlman PD, East ML. Divergent sapovirus strains and infection prevalence in wild carnivores in the Serengeti ecosystem: A long-term study. PLoS One. 2016; 11: 1–21. https://doi.org/10.1371/journal.pone.0163548 PMID: 27661997

69. Leary TP, Mushahwar IK. The GB Viruses. In: Mushawar IK, editor. Viral Hepatitis: Diagnosis, Therapy, and Prevention. Amsterdam: Elsevier; 2004. pp. 223–240.

70. Van der Kuyl AC. HIV infection and HERV expression: a review. Retrovirology. 2012; 9: 6. https://doi.org/10.1186/1742-4690-9-6 PMID: 22248111

71. Canuti M, van Beveren NJM, Jazaeri Farsani SM, de Vries M, Deijs M, Jebbink MF, et al. Viral metagenomics in drug-naïve, first-onset schizophrenia patients with prominent negative symptoms. Psychiatry Res. Elsevier; 2015; 229: 678–684. https://doi.org/10.1016/j.psychres.2015.08.025 PMID: 26304023

72. Somasekar S, Lee D, Rule J, Naccache S, Stone M, Busch MP, et al. Viral Surveillance in Serum Samples from Patients with Acute Liver Failure by Metagenomic Next-Generation Sequencing. Clin Infect Dis. 2017; IN PRESS. https://doi.org/10.1093/cid/cix596

73. Colson P, Fancello L, Gimenez G, Armougom F, Desnues C, Fournous G, et al. Evidence of the megavirome in humans. J Clin Virol. Elsevier B.V.; 2013; 57: 191–200. https://doi.org/10.1016/j.jcv.2013.03.018 PMID: 23664726

74. Rosseel T, Ozhelvaci O, Freimanis G, Van Borm S. Evaluation of convenient pretreatment protocols for RNA virus metagenomics in serum and tissue samples. J Virol Methods. Elsevier B.V.; 2015; 222: 72–80. https://doi.org/10.1016/j.jviromet.2015.05.010 PMID: 26025457

5 DISCUSIÓN

5.1 Eficiencia de la concentración de virus, bacterias y protozoos en el agua

La preparación, tratamiento y más específicamente la concentración de patógenos en muestras de agua son pasos críticos para una correcta cuantificación de la contaminación microbiana, el desarrollo de estudios de riesgo y la elaboración de regulaciones que protejan a la población. En este estudio se ha caracterizado el método de floculación con leche descremada para la concentración simultanea de virus (incluyendo virus con envolturas), bacterias y parásitos. Además, se correlacionaron los resultados de recuperación y se calculó la incertidumbre debida al método de concentración evaluado. También incluimos dentro de la evaluación la susceptibilidad de las partículas víricas a la floculación haciendo ensayos de infectividad.

Las recuperaciones más altas obtenidas en este experimento fueron con HAdV y *E. coli*. Este método de concentración ha sido utilizado anteriormente para evaluar muestras previamente inoculadas con HAdV, las cuales fueron cuantificadas mediante qPCR y con porcentajes de recuperación de 52% para agua de río, entre 41 y 50% para agua de mar y de entre 30 a 95% para agua residual (Calgua et al., 2013a, 2013b, 2008). El método de floculación con leche descremada también ha sido utilizado para otros virus como NoV con recuperaciones de entre 34 y 74% (Calgua et al., 2013a). En este estudio se describen los porcentajes de recuperación con sus respectivos intervalos de confianza, obtenidos a partir de un tamaño de muestra representativo lo que nos permite sugerir que el método de floculación con leche descremada junto con una cuantificación mediante qPCR pueden ser utilizados con la misma eficiencia en diferentes tipos de matrices de agua.

La floculación con leche descremada ha mostrado también ser muy útil en estudios de "microbial source tracking" (Rusiñol et al., 2015). Con este método se han detectado: JC poliomavirus, adenovirus porcinos y poliomavirus bovinos en muestras de aguas superficiales (Bofill-Mas et al., 2011); NoV, HAdV, JCPyV, RoV, Klassevirus, Asfavirus-like y poliomavirus de células de Merkel en aguas de río (Calgua et al., 2013a; Rusiñol et al., 2015, 2014); HAdV, RoV, el bacteriófago PP7 y NoV en agua de mar y HAdV, JCPyV y NoV en agua residual. Finalmente, el SMF ha sido modificado para detectar HAdV y NoV en fresas con muy buenos resultados (Melgaço et al., 2016).

Los virus con envoltura como el BVDV parecen ser más estables en el agua de lo que se esperaba. Teniendo en consideración la escasa información con la que se cuenta acerca de este tipo de virus, se decidió además de HAdV, RoV y el bacteriófago MS2, añadir el BVDV, un virus con envoltura lipídica que es patógeno para el ganado bovino, a las muestras de agua analizadas. El porcentaje de recuperación fue evaluado por triplicado y mediante los dos métodos de cuantificación disponibles: qPCR e infectividad. El BVDV se suele transmitir de una forma horizontal a través de la inhalación o la ingestión de partículas (MacLachlan and Dubovi, 2011). No existe información previa que describa que este virus puede sobrevivir en el agua, los resultados de este estudio sugieren que este virus podría transmitirse también a través del agua contaminada de los bebederos.

Dados los porcentajes de recuperación obtenidos por qPCR: 66% (HAdV 35), 24% (MS2), 28% (RoV SA-11), 15% (BVDV), 60% (*E. coli*), 30% (*H. pylori*) y 21% (*A. castellanii*) y los porcentajes de recuperación mediante infectividad: 59% (HAdV 35), 12% (MS2), 26% (RoV SA-11) y 0.7% (BVDV), podemos afirmar que el SMF es un método eficiente para concentrar microorganismos en el agua. La disponibilidad de técnicas efectivas y de bajo coste para la concentración simultanea de virus, bacterias y protozoos presentes en el agua es de gran importancia para el monitoreo de la calidad del agua (Hill et al., 2015). Además, estos métodos serán de gran ayuda para la aplicación de técnicas de secuenciación masiva que servirán para caracterizar la población microbiológica del agua. Hasta la fecha, no existe información disponible en la literatura de métodos de concentración que permitan la evaluación de volúmenes representativos de agua (10 litros), que puedan ser utilizados en diferentes matrices de agua, con un alto porcentaje de recuperación y que sirvan para evaluar virus, bacterias y protozoos simultáneamente.

Debido a la importancia de contar con un adecuado indicador de contaminación, se realizaron correlaciones de las recuperaciones entre los microorganismos evaluados. Teniendo en consideración los principios de la floculación, los flóculos adsorben las partículas presentes en el agua previamente acidificada, lo que incrementa el peso de las partículas facilitando su precipitación (Calgua et al., 2008). Los resultados observados en este estudio sugieren que este fenómeno es dependiente del tipo de microorganismo. Por este motivo, la utilización de un solo indicador de contaminación fecal parece no ser posible; sin embargo, la correlación encontrada entre bacterias por un lado y entre virus por otro, sugieren que tanto *E. coli* como HAdV son indicadores adecuados para determinar contaminación fecal bacteriana y viral respectivamente.

La relación entre los valores cuantificados por técnicas moleculares y por infectividad sugiere que esta diferencia sería debida a la proporción de partículas no infectivas y/o a la sensibilidad del método. Encontramos estudios en la literatura donde se observan cuantificaciones de HAdV en agua mediante qPCR entre 1 y 2 logaritmos más altos que mediante métodos basados en cultivo (Rames et al., 2016). HAdV y RoV son conocidos por ser resistentes al pH (Attoui et al., 2012; Harrach et al., 2012), por lo que su infectividad no se vio afectada por la acidificación de la muestra durante el proceso de SMF. El virus MS2 al ser más sensible al pH que los otros virus estudiados, muestra también diferencias tanto en la cuantificación como en los porcentajes de recuperación entre métodos. Por otro lado, el BVDV (nuestro virus modelo con envoltura) fue el que mostró las mayores diferencias entre métodos moleculares y de infectividad (ratio de 2.03), lo que mostraría una reducción de virus infectivos en el proceso de floculación, probablemente debido al pH ácido. Tradicionalmente los métodos de infectividad suelen dar cuantificaciones menores que las técnicas de qPCR (Amini and Kraatz, 2014).

Es importante tener en consideración que el modelo descrito en este estudio nos permite describir el carácter aleatorio de los porcentajes de recuperación desconocidos, pero no nos permite describir el carácter aleatorio de las concentraciones en el ambiente. Por lo tanto, el error ajustado en las expresiones matemáticas descritas en este artículo hace referencia al error técnico de medición única y exclusivamente.

La caracterización de la recuperación e incertidumbre con resultados de recuperación elevados y elevada reproducibilidad, apoyan la utilización de este método en muestras ambientales con un control de proceso como el bacteriófago MS2. La verificación de la recuperación con este virus debería oscilar entre el 10 y el 38% (correspondiente a 2 desviaciones estándares de la media de recuperación). Este intervalo nos permite describir la variabilidad e incertidumbre del método de SMF en nuestro laboratorio. Otra forma sería determinar el porcentaje de recuperación de éste u otro virus en cada laboratorio teniendo en cuenta que el virus utilizado no tiene que estar presente en la muestra con anterioridad. Aunque la variabilidad es una característica intrínseca de cada variable, la incertidumbre introduce una variabilidad “subjetiva” en la variable de interés (Vose, 2008). Esta última variabilidad, para nuestro caso en particular, puede incrementarse o disminuirse debido a múltiples factores como la experiencia del operario, el equipamiento y la calidad de los reactivos; lo que podría causar recuperaciones diferentes entre operarios o incluso laboratorios y deben ser tomados en cuenta.

Los valores de recuperación obtenidos en este estudio también son en alguna medida mejorables; es importante minimizar las pérdidas de los flóculos debido a la adhesión a las paredes laterales del cubo, después del paso de sedimentación o las perdidas debidas al proceso de decantación después del paso de centrifugación a 8000 rpm por 30 min. Los materiales fungibles utilizados por muestra en nuestro estudio son de bajo coste y pueden ser desinfectados y reutilizados para otros propósitos.

Los parámetros de variabilidad e incertidumbre descritos en este estudio para HAdV y *E. coli* serán de gran utilidad para futuras investigaciones. Bajo condiciones controladas, la variabilidad de la recuperación para HAdV (IC 95%: 53.5 – 78.5%) incluye los resultados de estudios previos que utilizaron este método (Calgua et al., 2008; Rusiñol et al., 2015, 2014), por este motivo, nosotros demostramos la importancia de modelar la recuperación con el fin de tener una mejor aproximación del riesgo real de la presencia de microorganismos en el agua.

La distribución de la recuperación bajo condiciones controladas no sugiere que la distribución de los microrganismos siga esa misma tendencia. Es importante tener en consideración que los métodos de muestreo ambiental siempre tienen limitaciones y requieren mucho trabajo para la estimación para describir la distribución de microorganismos en el ambiente (Petterson et al., 2015). La calidad microbiológica del agua puede variar mucho y muy rápidamente. Picos en la concentración de patógenos en cortos períodos de tiempo pueden incrementar el riesgo considerablemente y desencadenar brotes de enfermedades transmitidas por el agua, de tal manera que para cuando la contaminación es detectada, la mayoría de personas ya estuvieron expuestas (Gorchev and Ozolins, 2011).

Mientras que el riesgo se estima teniendo en cuenta la frecuencia de contaminación y la distribución de la dosis, la probabilidad de infección de un individuo en particular es estimada en función del número de microrganismos ingeridos (Ross, 2008). Los errores debido a la falta de precisión en la cuantificación casi siempre suelen infravalorar la concentración real. Por lo tanto, los métodos de extrapolación permiten una estimación realista de la concentración de microrganismos en las muestras de agua, la cual favorece a la obtención de mejores aproximaciones en los QMRA.

5.2 Análisis del riesgo microbiológico de la aplicación de aguas regeneradas en el riego de vegetales

En este estudio, los datos obtenidos de la concentración de dos virus (HAdV y NoV GII) provenientes de dos WWTPs, analizadas en dos proyectos europeos de nuestro grupo: Viroclime y Metawater, fueron utilizados para caracterizar la concentración del agua de entrada y la tratada por un período de 12 meses. Estas cuantificaciones se utilizaron para establecer una concentración adecuada que sea la base para un estudio de QMRA, tal como lo recomendada la WHO como método para la determinación del riesgo de consumir vegetales regados con agua regenerada (WHO, 2016a, 2006).

Debido a las dificultades asociadas con la detección de patógenos en agua, indicadores fecales bacterianos (FIB), en los que se incluyen coliformes fecales, *E. coli* y enterococos, son los que se utilizan para controlar la calidad en las WWTPs. FIB se comportan de manera muy diferente a los virus en el agua residual, siendo las bacterias más susceptibles a las condiciones ambientales y los tratamientos en las WWTPs (McMinn et al., 2017). Aunque los virus no son los únicos patógenos presentes en el agua que pueden afectar la salud, el riesgo de enfermedad es 10 a 10000 veces mayor para virus que para bacterias al mismo nivel de exposición (Haas et al., 1993). Por esta razón, la evaluación de únicamente FIB subestima el riesgo de virus entéricos para la salud de la población.

Uno de los principales problemas en los estudios de QMRA es la escasa información disponible para establecer una distribución que permita describir la concentración de microorganismos. Desafortunadamente, los métodos para cuantificar virus después de los procesos de tratamiento que se realizan en las WWTPs arrojan muchos valores negativos. Estos resultados se producen debido a que frecuentemente las concentraciones están por debajo de los límites de detección de las técnicas utilizadas y que éstas poseen una baja/moderada sensibilidad, especificidad y exactitud. Los resultados negativos obtenidos en muchos estudios rara vez demuestran una eliminación completa de los indicadores microbiológicos, como normalmente demuestra *E. coli* en los análisis rutinarios de las WWTPs (De Keuckelaere et al., 2015; Mok et al., 2014; Petterson and Ashbolt, 2016; Schijven et al., 2011; WHO, 2016a). Por estas razones, se han elaborado modelos matemáticos que se utilizan para hacer frente a estos problemas y obtener mejores aproximaciones de la concentración de microorganismos, teniendo en cuenta la concentración antes y después del tratamiento (Teunis et al., 1999, 2009). El monitoreo microbiológico antes y

después del proceso de tratamiento es la forma más directa de evaluar la eficacia del tratamiento (Smeets et al., 2010), lo que permite describir tanto las concentraciones bajas más probables como las ocasionadas en los casos de epidemias. Estos procedimientos son recomendados para análisis QMRA (WHO, 2016a).

En esta parte de la tesis se caracterizó el tratamiento en dos WWTPs con un tratamiento secundario convencional, pero con diferentes tratamientos terciarios. Para ello se utilizó el método no emparejado debido a lo complicado que significa seguir la muestra de agua a través del proceso entero del tratamiento. Los resultados sugieren que el tratamiento natural de la WWTP 2 (lagunaje natural) es en promedio más eficiente para reducir la carga viral, sin embargo, la gran variabilidad en la reducción de este sistema y el gran espacio necesario para tratar todo el volumen de agua que ingresa a la planta (actualmente solo tratan el 10% del volumen total), hace este proceso poco factible. El sistema más sofisticado de la WWTP 1 (Ultravioleta, cloración y Actiflo®) tiene una reducción menor, pero con una variabilidad más controlada en el proceso, claro está que vienen a expensas de un mayor gasto de energía.

La mayoría de parámetros utilizados en la evaluación de la exposición para la irrigación de vegetales usadas en esta tesis se han descrito previamente en anteriores estudios de QMRA publicados alrededor del mundo. Adicionalmente, nuestro estudio incluye algunas modificaciones que consideramos importantes a la hora de describir la dosis correcta como es la recuperación descrita para el método de SMF (primer capítulo) (Gonzales-Gustavson et al., 2017). Es bien conocido que las tasas de recuperación viral provenientes de la concentración y detección por métodos moleculares varia drásticamente, usualmente infravalorando la concentración real entre 1 a 3 órdenes de magnitud (Mok and Hamilton, 2014; Petterson et al., 2015). La eficiencia en la recuperación debe ser tomada en cuenta para cada método en particular utilizado para determinar la concentración y dentro de cada modelo. La ventaja de incluir a recuperación de manera estocástica es que este valor variara entre muestras dentro de la simulación.

Otro componente principal dentro del QMRA es el modelo de dosis-respuesta, el cual describe la relación entre la dosis y la probabilidad de un resultado, ya sea infección o enfermedad. Por tal motivo la elección de un modelo de dosis-respuesta puede afectar enormemente la determinación del riesgo. Existen en la actualidad varios modelos de dosis-respuesta (Van Abel et al., 2017), aunque muchos estudios siguen utilizando la aproximación a la Beta-Poisson y la exponencial como modelos para NoV y HAdV, respectivamente, nosotros elegimos los modelos recientemente

disponibles en la literatura en función de las nuevas investigaciones. Así tenemos, 1) la publicación reciente de un modelo de HAdV que tiene como ventaja describir la probabilidad de infección en función de la vía de inoculación (Teunis et al., 2016), el modelo anterior de dosis-respuesta (exponencial) utilizado para HAdV fue desarrollado con cepas causantes de enfermedad respiratoria, lo que limita su utilización en estudios de QMRA (Ashbolt, 2015); 2) La función hipergeométrica de dosis-respuesta para NoV, que además de describir mejor la probabilidad de infección, puede incluir el efecto de agregación de estos virus, fenómeno que se describe en la literatura particularmente en muestras de agua, característica que el modelo anterior de Beta-Poisson no es capaz de aportar (Teunis et al., 2008; Teunis and Havelaar, 2000). Sin embargo, el efecto de agregación no está completamente aceptado dentro de la comunidad científica principalmente porque se presume que las WWTPs deben ser capaces de eliminar con facilidad las partículas de mayor tamaño. Además, tampoco existe la seguridad de si la agregación se produce en aguas residuales y regeneradas. Una comparación entre el modelos demuestra que el que tiene en cuenta la agregación da como resultado una menor probabilidad de infección a dosis bajas comparado con el que no lo incluye, lo que podría infravalorar el riesgo (Mcbride, 2014; Van Abel et al., 2017).

Debido a la falta de consenso en la selección de parámetros adecuados para la evaluación de los efectos en la salud, en este estudio se desarrollaron en paralelo dos escenarios: uno conservador y otro desfavorable, estos van en función del modelo de dosis-respuesta y de la probabilidad de infección condicionada a enfermedad, los cuales están descritos en la literatura. El escenario desfavorable se ve representado en los valores más extremos de la distribución de los modelos en el escenario conservador. Casi en todos los casos, los resultados demuestran que el sistema falló en remover la suficiente cantidad de virus para alcanzar las recomendaciones establecidas por la WHO de 10^{-6} DALYs pppy. Como consecuencia, ambas WWTPs no fueron capaces de llegar al umbral de riesgo aceptable considerando el tratamiento terciario utilizado, lo que indica que la capacidad de reducción viral requiere la adición de más tratamientos. Únicamente en el escenario conservador de la WWTP 2 con NoV GII el riesgo fue inferior a las recomendaciones de la WHO.

Se tiene por establecido que los métodos moleculares como la q(RT)PCR sobreestiman el número de virus infectivos, especialmente después de procesos de desinfección que suelen inactivar las partículas víricas dejando el ácido nucleico intacto (Pecson et al., 2009; Shin and Sobsey, 2008). Sin embargo, estudios que evalúan la depuración biológica de fangos activos usados como

tratamiento secundario muestra que la reducción es debida principalmente a la eliminación física de partículas víricas, más que en la inactivación (Hata et al., 2013). En *E. coli*, las reducciones que incluyen el tratamiento de fangos activos, cloración y filtros de arena varían entre 3 a 5 logaritmos de reducción. Sin embargo, la reducción viral en el mismo proceso es de entre 1 y 3.5, lo que demuestra que las WWTPs no están preparadas para reducir eficientemente los virus del agua (Hata et al., 2013; Petterson and Stenström, 2015). las FIB son inactivadas mucho más rápidamente por la acción del cloro en comparación con algunos virus y bacterias (Ashbolt et al., 2001). Tratamientos adicionales como el Actiflo® describen reducciones de entre 1 a 3 logaritmos de colifagos bajo condiciones experimentales pero es importante reconocer que la efectividad de esta tecnología se ve comprometida por varios factores como la calidad del agua cruda y la sensibilidad de los microorganismos presentes al tratamiento (Mok et al., 2014).

Para la estimación de los DALYs en NoV GII se seleccionaron los parámetros utilizados previamente en otro estudio de QMRA en el área de estudio que incluye una estimación mixta de años de vida con discapacidad y años de vida perdidos con valores referenciales obtenidos en Cataluña y Holanda (Sales-Ortells et al., 2015), los resultados fueron comparados con los parámetros de carga de enfermedad por caso y fracción de susceptibilidad usados en otro estudio de Canadá (Chhipi-Shrestha et al., 2017), obteniendo similares resultados. Por lo tanto, debido a que no se cuenta con parámetros específicos para HAdV se decidió tomar en cuenta los parámetros canadienses para estimar los DALYs de este virus.

El análisis de sensibilidad preliminar mostró que la incertidumbre en la reducción de los microorganismos en las WWTPs y la concentración del agua de entrada son las variables de mayor influencia en la estimación del riesgo de consumo de vegetales irrigados con esta agua. Aunque la concentración tanto de NoV GII como de HAdV en el agua de entrada parece ser diferente entre las plantas, solo se pudo establecer diferencias estadísticas significativas en la concentración de HAdV.

Debido a que la evaluación rutinaria de virus en agua regenerada no está establecida, la presencia de virus entéricos humanos en agua es en cierta medida desconocida a menos que un brote sea reportado y usualmente cuando se colectan muestras, rara vez es posible identificar el origen (Gibson, 2014; Gorchev and Ozolins, 2011). Desafortunadamente, teniendo en cuenta los límites de detección para HAdV y NoV GII con SMF de 28.6 and 291 CG/100ml, es imposible determinar

con una única medición, que las WWTPs satisfacen los requerimientos de la WHO, dado que se requerirían analizar volúmenes muy grandes.

En este estudio se estimó el riesgo de salud asociado con el consumo de vegetales irrigados con efluente terciario proveniente de dos WWTPs y conteniendo NoV GII y HAdV, basándonos en un método que tiene en consideración las concentraciones de entrada y de salida e introduciendo el parámetro de la recuperación del método de concentración (SMF). Estos dan como resultado una estimación de la concentración de virus más realista después del tratamiento en la WWTP y en el laboratorio.

5.3 Aplicación de técnicas de metagenómica para la identificación de virus de transmisión hídrica y posibles causantes de hepatitis

El objetivo del tercer estudio en esta tesis fue investigar la presencia de virus en suero de pacientes previamente diagnosticados con hepatitis aguda, pero sin etiología identificada. Para tal propósito, las muestras de suero fueron agrupadas en diferentes pools para ser estudiadas por técnicas de metagenómica con la plataforma ilumina Mi-Seq. Los hallazgos demuestran la presencia de una gran variedad de secuencias virales en todos los pools compuestos por pacientes con hepatitis e y de etiología desconocida.

El virus de la hepatitis E fue detectado, además de en el pool de pacientes HEV+ utilizado como control positivo, también en el pool de pacientes autoinmunes/inmunosuprimidos (pool Ai+ImSP). En el pool de pacientes con HEV se identificaron contigs que alinearon con el genotipo HEV 3f. Este genotipo ha sido descrito previamente como un causante de brotes de hepatitis en Cataluña (Riveiro-Barciela et al., 2015), otras regiones de España (Rivero-Juarez et al., 2017) y en el sur de Francia (Legrand-Abravanel et al., 2009). Además también está relacionado con el consumo de productos derivados de carnes de cerdos y jabalíes y es considerada como una infección zoonótica transmitida a través de los alimentos principalmente (Banks et al., 2004; Meng et al., 1997; Riveiro-Barciela et al., 2015). En una evaluación detallada de las muestras individuales que conformaban el pool Ai+ImSP reanalizada con posterioridad mediante la técnica de PCR anidada, se observó que uno de los pacientes dio positivo al virus de la hepatitis E, el análisis de los contigs obtenidos en el mencionado pool corresponden a un genotipo diferente, el HEV 3a. La técnica de metagenómica tiene la ventaja de identificar la diversidad de virus dentro de los pools, lo que facilita la descripción de múltiples infecciones en una muestra única.

El protocolo bioinformático desarrollado en el laboratorio, principalmente por Josep Abril y Natalia Timoneda permite hacer un estudio detallado de las secuencias obtenidas. Se analizaron al menos tres tipos diferentes de contigs: a) contigs que alinean con secuencias previamente caracterizadas en una familia; b) contigs que están relacionadas a secuencias aún no clasificadas dentro de una familia, pero descritas con anterioridad y; c) contigs con una pequeña homología con secuencias previamente descritas, lo que correspondería a potenciales nuevos virus.

Con respecto a los anellovirus, el criterio de demarcación para establecer género en esta familia usa como punto de corte el 35% de identidad en la región del ORF1. Debido al número de cuasiespecies descubiertas en esta familia (Spandole et al., 2015), es difícil establecer un punto de corte para la determinación de especies.

En este estudio describimos la presencia de secuencias de anellovirus previamente asociadas a hepatitis como son los TTV-1, 11, 16 y SEN virus H (Luo et al., 2002; Okamoto, 2009); también identificamos virus de esta familia que han sido recientemente descritos en muestras de suero de pacientes infectados con HIV (P 19-3,P13-4,P9-6 y P1-3); o con otras diversas patologías tales como leucemia linfoblástica (TTV 10)(Chu et al., 2011), periodontitis gingival (TTMV 18) (Zhang et al., 2016), hemofilia (TTMDV MDJN47 y MDJN97) (Ninomiya et al., 2007) y finalmente en mujeres embarazadas cuyos hijos desarrollaron leucemia o linfomas (TTV S45, S57, S69, S72, S80 y S97) (Bzhalava et al., 2016).

El análisis de metagenómica ha permitido el descubrimiento de potenciales nuevas especies dentro de esta familia, Bzhalava et al 2016 describió por primera vez un grupo de secuencias en muestras de humanos, desvelando un nuevo grupo de anellovirus en humanos que no entraría dentro de ninguno de los géneros previamente descritos. En nuestro estudio, dos contigs (el 125 y el 1199) alinearon dentro de este potencial nuevo género de *Anelloviridae*, ambas con aproximadamente 70% de identidad con las descritas previamente en un estudio que evaluó muestras de suero de mujeres embarazadas (Bzhalava et al., 2016). Estos resultados sugieren que podría haber muchos más de estos virus dentro de esta familia esperando ser descubiertos.

Recientemente, el aislamiento de virus ADN de cadena simple (similares a los anellovirus) en muestras de suero y productos lácteos provenientes de bisontes europeos ha promovido investigaciones que podrían vincularlos con el desarrollo de enfermedades crónicas en humanos e incluso cáncer. Se sospecha de la actividad enzimática que podría tener algunas proteínas virales de estos sería la responsable de la producción de glucósidos tóxicos y utilizando como sustrato los

glucósidos de productos cárnicos y lácteos. Estos glucósidos tendrían un actividad inmunogénica causando reacciones autoinmunes exacerbadas que se observan comúnmente en diversas patologías crónicas, entre ellas esclerosos múltiple, o en diversos tipos de cáncer como el de colon, senos y próstata (Zur Hausen et al., 2017).

La presencia de anellovirus en este estudio no necesariamente indicaría que estos virus son necesariamente los agentes causales de las hepatitis agudas. Estudios previos han sugerido que la presencia de TTV en plasma sanguíneo podría ser indicador del estado de salud del sistema inmune (Touinssi et al., 2001). Otro estudio demuestra que pacientes inmunocomprometidos o que se encuentran recibiendo una terapia inmunosupresiva presentan un incremento en la carga de anellovirus en plasma (Li et al., 2013). La secuenciación masiva de muestras de plasma en un estudio realizado durante varios meses en pacientes que recibieron algún tipo de trasplante muestra también una carga viral incrementada, mientras que la población bacteriana permanece estable durante el mismo tiempo de observación (De Vlaminck et al., 2013; Hofer, 2014).

Nuestros resultados también describen la presencia de una cepa particular de sapovirus en todos los pools de pacientes con hepatitis clínica de etiología desconocida. Este virus esta clasificado dentro del genogrupo 5, cepa 2 (GV.2) y ha sido recientemente descrito en muestra de heces de humanos procedentes en un brote de gastroenteritis en Japón, posiblemente atribuida al consumo de alimentos y utilizando el método de secuenciación masiva (Shibata et al., 2015). Fragmentos parciales de este virus fueron descritos con anterioridad en otros brotes de gastroenteritis en Italia (Medici et al., 2012), en muestras de agua del rio Llobregat en Barcelona (la misma región donde se realizó este estudio) (Sano et al., 2011) y en aguas residuales del Japón (Hansman et al., 2007). Todas estas evidencias sugieren que este virus es altamente prevalente en diferentes áreas del planeta.

Algunos contigs encontrados en este estudio también alinearon con un miembro de la familia *Astroviridae.* El astrovirus VA3 fue identificado en la mayoría de los pools conformados por pacientes con hepatitis de etiología desconocida. Sin embargo, estos contigs se encontraron en menor cantidad y fueron de menor longitud con respecto a los que alinearon con sapovirus. La primera descripción de este virus en particular fue en heces de pacientes pediátricos con signos de diarrea en la India (Finkbeiner et al., 2009), y posteriormente secuenciado completamente (Jiang et al., 2013). Este virus también ha sido descrito en muestras de heces de la zona sur de China (Xiao et al., 2011), Kenia y Gambia (Meyer et al., 2015). Al igual que con sapovirus, no existen

reportes previos que asocien este virus con hepatitis, sin embargo, existen sapovirus que son reconocidos como causantes de hepatitis en patos (Liu et al., 2014) y otros que han sido aislados de hígado de ratones (Yokoyama et al., 2012). El rol del sapovirus VA3 en el desarrollo de alguna enfermedad permanece desconocido.

La presencia de virus causantes de gastroenteritis en muestras de suero ha sido descrita con anterioridad. El NoV ha sido identificado en muestras de sangre de niños con gastroenteritis aguda, en adultos inmunocomprometidos y en cerdos y bovinos gnotobióticos (Cheetham et al., 2006; Frange et al., 2012; Fumian et al., 2013; Lemes et al., 2014; Souza et al., 2008; Takanashi et al., 2009) lo que sugiere que la infección con NoV puede no estar limitada a los intestinos. En contraposición, otro estudio menciona que no ha sido posible detectar NoV en muestras de adultos inmunocomprometidos (Newman et al., 2015). El rol patogénico de Sapovirus GV.2 y de astrovirus VA3 detectados en el suero de estos pacientes es desconocido, aunque nuestros resultados sugieren que la presencia de estos virus en los pools de pacientes con hepatitis de etiología desconocida, incluyendo el pool Ai+ImSp, requiere más investigación, ya que no existe evidencia previa que permita relacionar los mencionados virus con hepatitis en humanos.

Finalmente, los retrovirus endógenos humanos son remanentes de retrovirus integrados dentro del genoma humano y son considerados funcionalmente defectivos (Van der Kuyl, 2012). Estos virus son reportados frecuentemente en estudios de metagenómica en altas concentraciones (Li et al., 2013; Van der Kuyl, 2012) sin estar asociados con ninguna patología en particular (Canuti et al., 2015). Los resultados obtenidos en este estudio demuestran que su presencia no estaría asociada a la enfermedad en estudio.

El análisis de metagenómica aplicado en esta tesis permitió detectar un amplio espectro de especies virales basadas en secuencias obtenidas en los pools evaluados. Estas incluyen diferentes genotipos de HEV, una gran diversidad de anellovirus; incluyendo secuencias de virus potencialmente nuevos y finalmente el sapovirus GV.2 y el astrovirus VA3, ambos de descripción reciente y que han sido relacionados con gastroenteritis. El rol de estos virus como posibles agentes causales de hepatitis queda abierto a futuros estudios. El análisis de metagenómica ofrece un gran abanico de posibilidades en el diagnóstico, caracterización e identificación de nuevos virus y coinfecciones, las cuales serán de gran importancia para la comprensión de patologías con etiología desconocida hasta el momento.

6 CONCLUSIONES

Los objetivos desarrollados en la presente tesis doctoral han dado lugar a una serie de resultados publicados o en proceso de publicación; las principales conclusiones de estos trabajos se detallan a continuación:

- El método SMF presenta un buen nivel de eficiencia y repetitividad para la concentración simultánea de virus, bacterias y protozoos.
- Los métodos de cuantificación por q(RT)PCR e infectividad evaluados muestran que los porcentajes de recuperación son similares para los virus HAdV y RoV; por lo que la concentración viral mediante el método de SMF permitiría cuantificar estos virus por técnicas moleculares e infectividad.
- La cuantificación del BVDB mediante el método de infectividad después de la concentración se vio afectada por la sensibilidad de este virus al pH y la cuantificación de este virus se pudo realizar de manera más eficiente a través de técnicas moleculares de q(RT)PCR.
- La estimación de la concentración de un microorganismo usando el método de SMF y q(RT)PCR con la predicción de intervalos al 95% mediante “bootstrap” de las recuperaciones, permitió la obtención de intervalos de predicción de la concentración real de los patógenos e indicadores. Este método puede ser aplicado como medida de incertidumbre en estudios de QMRA que utilicen el método de concentración de SMF.
- Se desarrolló un modelo para cuantificar el riesgo de salud asociado al consumo de lechugas irrigadas con agua regenerada proveniente de dos WWTPs con tratamientos terciarios diferentes. Se aplicó un modelo estocástico de QMRA para estimar los DALYs correspondientes para HAdV y NoV GII.
- Se encontraron diferencias entre la capacidad de reducción viral entre las WWTPs siendo el tratamiento terciario con lagunaje natural es más eficaz en promedio, pero con más variabilidad que la planta que utiliza UV, cloro y Actiflo®.
- Ninguna de las dos plantas evaluadas consiguió llegar al valor promedio en el riesgo de $<10^{-6}$ DALYs pppy considerado como aceptable por la WHO para irrigar vegetales con efluente terciario, a excepción de en la WWTP 2 con NoV GII en el escenario conservador.

- Esta es la primera descripción de un ensayo de QMRA que utiliza HAdV para evaluar el riesgo de consumo de vegetales contaminados con agua regenerada, y demuestra que se observan resultados similares a los obtenidos con NoV GII, tanto en el escenario desfavorable como en el conservador.
- La técnica de secuenciación masiva aplicada en esta tesis permitió detectar una gran variedad de secuencias de especies virales en las muestras de sueros de pacientes con hepatitis y las principales familias identificadas fueron *Anelloviridae*, *Caliciviridae* y *Astroviridae*, todas estas relacionadas con transmisión fecal-oral.
- Dentro de la familia *Anelloviridae* se describieron múltiples especies, predominantemente en pacientes con signos de hepatitis. Dentro de las secuencias encontradas se incluyen potenciales nuevas especies y algunas de ellas podrían ser parte de un nuevo género dentro de esta familia.
- El sapovirus GV.2 y el astrovirus VA3, dos virus de descripción reciente y asociados a gastroenteritis fueron identificados exclusivamente en los pools de pacientes con hepatitis aguda de etiología desconocida.
- La reproducibilidad entre las réplicas de pools de voluntarios sanos sugiere que la metodología aplicada de secuenciación masiva es un método de análisis robusto y fiable.
- El análisis de metagenómica aplicado en esta tesis nos ha permitido caracterizar la población viral de pacientes con hepatitis de etiología desconocida, lo que facilitaría la detección de coinfecciones y de posibles nuevas que deberían confirmarse en futuros estudios.

7 REFERENCIAS

Albinana-Gimenez, N., Clemente-Casares, P., Calgua, B., Huguet, J.M., Courtois, S., Girones, R., 2009. Comparison of methods for concentrating human adenoviruses, polyomavirus JC and noroviruses in source waters and drinking water using quantitative PCR. J. Virol. Methods 158, 104–109. https://doi.org/10.1016/j.jviromet.2009.02.004

Amini, K., Kraatz, H.-B., 2014. Recent advances and developments in monitoring biological agents in water samples. Rev. Environ. Sci. Bio/Technology 14, 23–48. https://doi.org/10.1007/s11157-014-9351-5

Ashbolt, N., Grabow, W., Snozzi, M., 2001. Indicators of microbial water quality, in: Fewtrell, L., Bartram, J. (Eds.), Water Quality: Guidelines, Standards and Health. IWA Publishing, London, pp. 289–316. https://doi.org/10.4324/9781315693606

Ashbolt, N.J., 2015. Microbial Contamination of Drinking Water and Human Health from Community Water Systems. Curr. Environ. Heal. reports 2, 95–106. https://doi.org/10.1007/s40572-014-0037-5

Attoui, H., Mertens, P., Becnel, J., Belaganahali, S., Betgoin, M., Brussaard, C., Chappell, J., Ciarlet, M., de Vas, M., Dermody, T., Dormitzer, P., Duncan, R., Fang, Q., Graham, R., Guglielmi, K., Harding, R., Hilman, B., Makkay, S., Marzachì, C., Matthijnssens, J., Milne, R., Mohd Jaafar, F., Mori, H., Noordeloos, A., Omura, T., Patton, J., Rao, S., Mann, M., Stolts, D., Suzuki, N., Upadhyaya, N., Wei, C., Zhou, H., 2012. Family reoviridae, in: King, A., Adams, M., Carstens, E., Lefkowitz, E. (Eds.), Virus Taxonomy, Clasification and Nomenclature of Viruses. ElSevier Academic Press, Davis, pp. 541–637.

Banks, M., Bendall, R., Grierson, S., Heath, G., Mitchell, J., Dalton, H., 2004. Human and porcine hepatitis E virus strains, United Kingdom. Emerg. Infect. Dis. 10, 953–955. https://doi.org/10.3201/eid1005.030908

Biagini, P., Bendinelli, M., Hino, S., Kakkola, L., Mankertz, A., Niel, C., Okamoto, H., Raidal, S., Teo, C., Todd, D., 2012. Anelloviridae, in: King, A., Adams, M., Carstens, E., Lefkowitz, E. (Eds.), Virus Taxonomy, Clasification and Nomenclature of Viruses. ElSevier Academic Press, Davis, pp. 331–341.

Bofill-Mas, S., Hundesa, A., Calgua, B., Rusinol, M., Maluquer de Motes, C., Girones, R., 2011. Cost-effective method for microbial source tracking using specific human and animal viruses. J. Vis. Exp. https://doi.org/10.3791/2820

Bofill-Mas, S., Rusiñol, M., Fernandez-cassi, X., Carratalà, A., Hundesa, A., Girones, R., 2013. Quantification of Human and Animal Viruses to Environmental Samples. Biomed Res. Int. 2013. https://doi.org/10.1155/2014/675857

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

Bostan, N., 2013. Current and Future Prospects of Torque Teno Virus. J. Vaccines Vaccin. 1–9. https://doi.org/10.4172/2157-7560.S1-004

Brinkman, N.E., Haffler, T.D., Cashdollar, J.L., Rhodes, E.R., 2013. Evaluation of methods using celite to concentrate norovirus, adenovirus and enterovirus from wastewater. J. Virol. Methods 193, 140–146. https://doi.org/10.1016/j.jviromet.2013.05.014

Bzhalava, D., Hultin, E., Muhr, L.S.A., Ekstrom, J., Lehtinen, M., Villiers, E.M. de, Dillner, J., 2016. Viremia during pregnancy and risk of childhood leukemia and lymphomas in the offspring: nested case-control study. Int. J. Cancer 138, 2212–2220. https://doi.org/10.1002/ijc.29666

Calgua, B., Fumian, T., Rusinol, 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. https://doi.org/10.1016/j.watres.2013.02.043

Calgua, B., Fumian, T., Rusiñol, M., Rodriguez-Manzano, J., Mbayed, V.A., Bofill-Mas, S., Miagostovich, M., Girones, R., 2013b. Detection and quantification of classic and emerging viruses by skimmed-milk flocculation and PCR in river water from two geographical areas. Water Res. https://doi.org/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., Lopez-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. https://doi.org/10.1016/j.jviromet.2008.08.003

Calgua, B., Rodriguez-Manzano, J., Hundesa, A., Sunen, 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, 215–221. https://doi.org/10.1016/j.jviromet.2012.10.012

Canuti, M., van Beveren, N.J.M., Jazaeri Farsani, S.M., de Vries, M., Deijs, M., Jebbink, M.F., Zaaijer, H.L., van Schaik, B.D.C., van Kampen, A.H.C., van der Kuyl, A.C., de Haan, L., Storosum, J.G., van der Hoek, L., 2015. Viral metagenomics in drug-naïve, first-onset schizophrenia patients with prominent negative symptoms. Psychiatry Res. 229, 678–684. https://doi.org/10.1016/j.psychres.2015.08.025

Cheetham, S., Souza, M., Meulia, T., Grimes, S., Han, M.G., Saif, L.J., 2006. Pathogenesis of a genogroup II human norovirus in gnotobiotic pigs. J. Virol. 80, 10372–81. https://doi.org/10.1128/JVI.00809-06

Chhipi-Shrestha, G., Hewage, K., Sadiq, R., 2017. Microbial quality of reclaimed water for urban reuses: Probabilistic risk-based investigation and recommendations. Sci. Total Environ. 576, 738–751. https://doi.org/10.1016/j.scitotenv.2016.10.105

Chivero, E.T., Stapleton, J.T., 2015. Tropism of human pegivirus (Formerly known as GB virus C/hepatitis G virus) and host immunomodulation: Insights into a highly successful viral infection. J. Gen. Virol. 96, 1521–1532. https://doi.org/10.1099/vir.0.000086

Chu, C., Zhang, L., Dhayalan, A., Agagnina, B.M., Magli, A.R., Fraher, G., Didier, S., Johnson, L., Kennedy, W.J., Damle, R.N., Yan, X.-J., Patten, P.E.M., Teichberg, S., Koduru, P., Kolitz, J.E., Allen, S.L., Rai, K.R., Chiorazzi, N., 2011. Torque Teno Virus 10 Isolated by Genome Amplification Techniques from a Patient with Concomitant Chronic Lymphocytic Leukemia and Polycythemia Vera. Mol. Med. 17, 1. https://doi.org/10.2119/molmed.2010.00110

De Keuckelaere, A., Jacxsens, L., Amoah, P., Medema, G., Mcclure, P., Jaykus, L.A., Uyttendaele, M., 2015. Zero Risk Does Not Exist: Lessons Learned from Microbial Risk Assessment Related to Use of Water and Safety of Fresh Produce. Compr. Rev. Food Sci. Food Saf. 14, 387–410. https://doi.org/10.1111/1541-4337.12140

De Vlaminck, I., Khush, K.K., Strehl, C., Kohli, B., Luikart, H., Neff, N.F., Okamoto, J., Snyder, T.M., Cornfield, D.N., Nicolls, M.R., Weill, D., Bernstein, D., Valantine, H.A., Quake, S.R., 2013. Temporal response of the human virome to immunosuppression and antiviral therapy. Cell 155, 1178–1187. https://doi.org/10.1016/j.cell.2013.10.034

Delwart, E., 2007. Viral metagenomics. Rev. Med. Virol. 17, 115–131. https://doi.org/10.1002/rmv.532

Estes, M., Greenberg, H., 2013. Rotaviruses, in: Knipe, D., Howley, P. (Eds.), Fields Virology. Lippincott Williams & Wilkins, Philadelphia, pp. 1347–1401.

Eusebi, L.H., Zagari, R.M., Bazzoli, F., 2014. Epidemiology of Helicobacter pylori Infection. Helicobacter 19, 1–5. https://doi.org/10.1111/hel.12165

Fernandez-Cassi, X., 2017. Aplicació de técniques de seqüenciació massiva a léstudi de virus potencialment contaminants dàiguas i/o aliments. Universitat de Barcelona.

Finkbeiner, S., Holtz, L., Jiang, Y., Rajendran, P., Franz, C., Zhao, G., Kang, G., Wang, D., 2009. Human stool contains a previously unrecognized diversity of novel astroviruses. Virol. J. 6, 161. https://doi.org/10.1186/1743-422X-6-161

Frange, P., Touzot, F., Debré, M., Héritier, S., Leruez-Ville, M., Cros, G., Rouzioux, C., Blanche, S., Fischer, A., Avettand-Fenoël, V., 2012. Prevalence and clinical impact of norovirus fecal shedding in children with inherited immune deficiencies. J. Infect. Dis. 206, 1269–1274. https://doi.org/10.1093/infdis/jis498

Frey, H.C., Patil, S.R., 2002. Identification and review of sensitivity analysis methods. Risk Anal. 22, 553–578. https://doi.org/10.1111/0272-4332.00039

Fumian, T.M., Justino, M.C.A., Mascarenhas, J.D.A.P., Reymão, T.K.A., Abreu, E., Soares, L., Linhares, A.C., Gabbay, Y.B., 2013. Quantitative and molecular analysis of noroviruses RNA in blood from children hospitalized for acute gastroenteritis in Belém, Brazil. J. Clin. Virol. 58, 31–35. https://doi.org/10.1016/j.jcv.2013.06.043

Gascón, J., 2006. Epidemiology, etiology and pathophysiology of traveler's diarrhea. Digestion 73, 102–108. https://doi.org/10.1159/000089785

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. https://doi.org/10.1016/j.watres.2016.11.020

Gibson, K.E., 2014. Viral pathogens in water: Occurrence, public health impact, and available control strategies. Curr. Opin. Virol. 4, 50–57. https://doi.org/10.1016/j.coviro.2013.12.005

Gonzales-Gustavson, E., Cárdenas-Youngs, Y., Calvo, M., da Silva, M.F.M., Hundesa, A., Amorós, I., Moreno, Y., Moreno-Mesonero, L., Rosell, R., Ganges, L., Araujo, R., Girones, R., 2017. Characterization of the efficiency and uncertainty of skimmed milk flocculation for the simultaneous concentration and quantification of water-borne viruses, bacteria and protozoa. J. Microbiol. Methods 134, 46–53. https://doi.org/10.1016/j.mimet.2017.01.006

Gorchev, H.G., Ozolins, G., 2011. WHO guidelines for drinking-water quality. WHO Chron. 38, 104–108. https://doi.org/10.1016/S1462-0758(00)00006-6

Griffin, J.S., Plummer, J.D., Long, S.C., 2008. Torque teno virus: an improved indicator for viral pathogens in drinking waters. Virol. J. 5, 112. https://doi.org/10.1186/1743-422X-5-112

Haas, C.N., Rose, J.B., Gerba, C., Regli, S., 1993. Risk assessment of virus in drinking water. Risk Anal. 13, 545–552.

Haas, C.N., Rose, J.B., Gerba, C.P., 1999. Quantitative Microbial Risk Assessment, First edit. ed. John Wiley & Sons, Inc., New York.

Hansman, G.S., Sano, D., Ueki, Y., Imai, T., Oka, T., Katayama, K., Takeda, N., Omura, T., 2007. Sapovirus in water, Japan. Emerg. Infect. Dis. 13, 133–135. https://doi.org/10.3201/eid1301.061047

Harrach, B., Benkö, M., Both, G., Brown, M., Davidson, A., Echavarría, M., Hess, M., Jones, M., Kajon, A., Lehmkuhl, H., Mautner, V., Mittal, S., Wadell, G., 2012. Family adenoviridae, in: King, A., Adams, M., Carstens, E., Lefkowitz, E. (Eds.), Virus Taxonomy, Clasification and Nomenclature of Viruses. ElSevier Academic Press, San Diego, pp. 125–141.

Hata, A., Kitajima, M., Katayama, H., 2013. Occurrence and reduction of human viruses, F-specific RNA coliphage genogroups and microbial indicators at a full-scale wastewater treatment plant in Japan. J. Appl. Microbiol. 114, 545–554. https://doi.org/10.1111/jam.12051

Havelaar, A.H., Melse, J.M., 2003. Quantifying public health risk in the WHO Guidelines for Drinking Water Quality: a burden of disease approach. Report 734301022/2003. 1–49.

Hill, V.R., Narayanan, J., Gallen, R.R., Ferdinand, K.L., Cromeans, T., Vinjé, J., 2015. Development of a nucleic Acid extraction procedure for simultaneous recovery of DNA and RNA from diverse microbes in water. Pathogens 4, 335–354. https://doi.org/10.3390/pathogens4020335

Hofer, U., 2014. Microbiome: anelloviridae go viral. Nat. Rev. Microbiol. 12, 4–5. https://doi.org/10.1038/nrmicro3192

Hsiao, K.L., Wang, L.Y., Lin, C.L., Liu, H.F., 2016. New phylogenetic groups of torque teno virus identified in eastern Taiwan indigenes. PLoS One 11, 1–10. https://doi.org/10.1371/journal.pone.0149901

Hugenholtz, P., Tyson, G.W., 2008. Metagenomics. Nature 455, 481–483. https://doi.org/10.1038/455481a

Jiang, H., Holtz, L.R., Bauer, I., Franz, C.J., Zhao, G., Bodhidatta, L., Shrestha, S.K., Kang, G., Wang, D., 2013. Comparison of novel MLB-clade, VA-clade and classic human astroviruses highlights constrained evolution of the classic human astrovirus nonstructural genes. Virology 436, 8–14. https://doi.org/10.1016/j.virol.2012.09.040

Kakkola, L., Bondén, H., Hedman, L., Kivi, N., Moisala, S., Julin, J., Ylä-Liedenpohja, J., Miettinen, S., Kantola, K., Hedman, K., Söderlund-Venermo, M., 2008. Expression of all six human Torque teno virus (TTV) proteins in bacteria and in insect cells, and analysis of their IgG responses. Virology 382, 182–189. https://doi.org/10.1016/j.virol.2008.09.012

Karavarsamis, N., Hamilton, A.J., 2010. Estimators of annual probability of infection for quantitative microbial risk assessment. J. Water Health 8, 365–373. https://doi.org/10.2166/wh.2010.045

Kundu, A., McBride, G., Wuertz, S., 2013. Adenovirus-associated health risks for recreational activities in a multi-use coastal watershed based on site-specific quantitative microbial risk assessment. Water Res. 47, 6309–6325. https://doi.org/10.1016/j.watres.2013.08.002

Lambertini, E., Spencer, S.K., Bertz, P.D., Loge, F.J., Kieke, B.A., Borchardt, M.A., 2008. Concentration of enteroviruses, adenoviruses, and noroviruses from drinking water by use of glass wool filters. Appl. Environ. Microbiol. 74, 2990–2996. https://doi.org/10.1128/AEM.02246-07

Leary, T.P., Mushahwar, I.K., 2004. The GB Viruses, in: Mushawar, I.K. (Ed.), Viral Hepatitis: Diagnosis, Therapy, and Prevention. Elsevier, Amsterdam, pp. 223–240.

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, 110–114.

https://doi.org/10.3201/eid1501.080296

Lemes, L.G.N., Corrêa, T.S., Fiaccadori, F.S., Cardoso, D. das Dô. de P., Arantes, A. de M., Souza, K.M.C., Souza, M., 2014. Prospective study on Norovirus infection among allogeneic stem cell transplant recipients: Prolonged viral excretion and viral RNA in the blood. J. Clin. Virol. 61, 329–333. https://doi.org/10.1016/j.jcv.2014.08.004

Li, L., Deng, X., Linsuwanon, P., Bangsberg, D., Bwana, M.B., Hunt, P., Martin, J.N., Deeks, S.G., Delwart, E., 2013. AIDS alters the commensal plasma virome. J. Virol. 87, 10912–5. https://doi.org/10.1128/JVI.01839-13

Liu, N., Wang, F., Shi, J., Zheng, L., Wang, X., Zhang, D., 2014. Molecular characterization of a duck hepatitis virus 3-like astrovirus. Vet. Microbiol. 170, 39–47. https://doi.org/10.1016/j.vetmic.2014.01.026

Luo, K., He, H., Liu, Z., Liu, D., Xiao, H., Jiang, X., Liang, W., Zhang, L., 2002. Novel variants related to TT virus distributed widely in China. J. Med. Virol. 67, 118–126. https://doi.org/10.1002/jmv.2200

MacLachlan, J.., Dubovi, E., 2011. Fenner´s Veterinary Virology, 4th ed. Academic Press, San Diego. https://doi.org/http://dx.doi.org/10.1016/B978-0-12-375158-4.00030-4

Madigan, M., Martinko, J., Bender, K., Buckley, D., Stahl, D., 2014. Brock Biology of Microorganisms, 14th ed. Pearson, Harlow.

Marciano-Cabral, F., Cabral, G., 2003. Acanthamoeba spp. as agents of disease in humans. Clin. Microbiol. Rev. 16, 273–307. https://doi.org/10.1128/CMR.16.2.273-307.2003

Mcbride, G.B., 2014. Norovirus dose-response in sewage-related QMRA: The importance of virus aggregation. Congr. Env. Model. Softw. Paper 69.

McMinn, B.R., Ashbolt, N.J., Korajkic, A., 2017. Bacteriophages as indicators of faecal pollution and enteric virus removal. Lett. Appl. Microbiol. 65, 11–26. https://doi.org/10.1111/lam.12736

Medici, M.C., Tummolo, F., Albonetti, V., Abelli, L.A., Chezzi, C., Calderaro, A., 2012. Molecular detection and epidemiology of astrovirus, bocavirus, and sapovirus in Italian children admitted to hospital with acute gastroenteritis, 2008-2009. J. Med. Virol. 84, 643–650. https://doi.org/10.1002/jmv.23231

Melgaço, F.G., Victoria, M., Correa, A.A., Ganime, A.C., Malta, F.C., Brandao, M.L.L., de Mello Medeiros, V., de Oliveira Rosas, C., Bricio, S.M.L., Miagostovich, M.P., 2016. Virus recovering from strawberries: Evaluation of a skimmed milk organic flocculation method for assessment of microbiological contamination. Int. J. Food Microbiol. 217, 14–19. https://doi.org/10.1016/j.ijfoodmicro.2015.10.005

Meng, X.J., Purcell, R.H., Halbur, P.G., Lehman, J.R., Webb, D.M., Tsareva, T.S., Haynes, J.S., Thacker, B.J., Emerson, S.U., 1997. A novel virus in swine is closely related to the human hepatitis E virus. Proc Natl Acad Sci U S A 94, 9860–9865.

Meyer, C.T., Bauer, I.K., Antonio, M., Adeyemi, M., Saha, D., Oundo, J.O., Ochieng, J.B., Omore, R., Stine, O.C., Wang, D., Holtz, L.R., 2015. Prevalence of classic, MLB-clade and VA-clade Astroviruses in Kenya and The Gambia. Virol. J. 12, 78. https://doi.org/10.1186/s12985-015-0299-z

Mi, Z., Yuan, X., Pei, G., Wang, W., An, X., Zhang, Z., Huang, Y., Peng, F., Li, S., Bai, C., Tong, Y., 2014. High-throughput sequencing exclusively identified a novel Torque teno virus genotype in serum of a patient with fatal fever. Virol. Sin. 29, 112–118. https://doi.org/10.1007/s12250-014-3424-z

Michen, B., Graule, T., 2010. Isoelectric points of viruses. J. Appl. Microbiol. 109, 388–397. https://doi.org/10.1111/j.1365-2672.2010.04663.x

Mok, H.-F., Hamilton, A.J., 2014. Exposure Factors for Wastewater-Irrigated Asian Vegetables and a Probabilistic Rotavirus Disease Burden Model for Their Consumption. Risk Anal. 34, 602–613. https://doi.org/10.1111/risa.12178

Mok, H.F., Barker, S.F., Hamilton, A.J., 2014. A probabilistic quantitative microbial risk assessment model of norovirus disease burden from wastewater irrigation of vegetables in Shepparton, Australia. Water Res. 54, 347–362. https://doi.org/10.1016/j.watres.2014.01.060

Newman, K.L., Marsh, Z., Kirby, A.E., Moe, C.L., Leon, J.S., 2015. Immunocompetent Adults from Human Norovirus Challenge Studies do not Exhibit Norovirus Viremia. J. Virol. 89, JVI.00392-15. https://doi.org/10.1128/JVI.00392-15

Ninomiya, M., Nishizawa, T., Takahashi, M., Lorenzo, F.R., Shimosegawa, T., Okamoto, H., 2007. Identification and genomic characterization of a novel human torque teno virus of 3.2 kb. J.

Gen. Virol. 88, 1939–1944. https://doi.org/10.1099/vir.0.82895-0

Nishizawa, T., Okamoto, H., Konishi, K., Yoshizawa, H., Miyakawa, Y., Mayumi, M., 1997. A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology. Biochem. Biophys. Res. Commun. 241, 92–7. https://doi.org/10.1006/bbrc.1997.7765

Ogilvie, L., Jones, B., 2015. The human gut virome: a multifaceted majority. Front. Microbiol. 6, 918. https://doi.org/10.3389/fmicb.2015.00918

Ohbayashi, H., Tanaka, Y., Ohoka, S., Chinzei, R., Kakinuma, S., Goto, M., Watanabe, M., Marumo, F., Sato, C., 2001. TT virus is shown in the liver by in situ hybridization with a PCR-generated probe from the serum TTV-DNA. J. Gastroenterol. Hepatol. 16, 424–428. https://doi.org/10.1046/j.1440-1746.2001.02460.x

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

Okamoto, H., 2009. History of discoveries and pathogenicity of TT viruses, in: de Villiers, E.-M., zur Hausen, H. (Eds.), TT Viruses - The Still Elusive Human Pathogens. Springer, Berlin, pp. 1–20. https://doi.org/10.1007/978-3-540-70972-5_1

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, 5544–5554. https://doi.org/10.1128/AEM.00425-09

Petterson, S., Grøndahl-Rosado, R., Nilsen, V., Myrmel, M., Robertson, L.J., 2015. Variability in the recovery of a virus concentration procedure in water: Implications for QMRA. Water Res. 87, 79–86. https://doi.org/10.1016/j.watres.2015.09.006

Petterson, S.R., Ashbolt, N.J., 2016. QMRA and water safety management: Review of application in drinking water systems. J. Water Health 14, 571–589. https://doi.org/10.2166/wh.2016.262

Petterson, S.R., Ashbolt, N.J., Sharma, A., 2001. Microbial risks from wastewater irrigation of salad crops: A screening-level risk assessment. Water Environ. Res. 73, 667–672. https://doi.org/10.2175/106143001x143402

Petterson, S.R., Stenström, T.A., 2015. Quantification of pathogen inactivation efficacy by free chlorine disinfection of drinking water for QMRA. J. Water Health 13, 625–644. https://doi.org/10.2166/wh.2015.193

Pevsner, J., 2015. Bioinformatics and functional genomics, Third Edit. ed. Wiley-Blackwell.

Pina, S., Puig, M., Lucena, F., Jofre, J., Girones, R., 1998. Viral pollution in the environment and in shellfish: Human adenovirus detection by PCR as an index of human viruses. Appl. Environ. Microbiol. 64, 3376–3382.

Prata, C., Ribeiro, A., Cunha, Â., Gomes, N.C.M., Almeida, A., 2012. Ultracentrifugation as a direct method to concentrate viruses in environmental waters: virus-like particle enumeration as a new approach to determine the efficiency of recovery. J. Environ. Monit. 14, 64–70. https://doi.org/10.1039/C1EM10603A

Previsani, N., Lavanchy, D., Siegl, G., 2004. Hepatitis A, in: Mushahwar Isa K (Ed.), Viral Hepatitis: Molecular Biology, Diagnosis, Epidemiology and Control. ElSevier, Amsterdam, pp. 1–30.

Rames, E., Roiko, A., Stratton, H., Macdonald, J., 2016. Technical aspects of using human adenovirus as a viral water quality indicator. Water Res. 96, 308–326. https://doi.org/10.1016/j.watres.2016.03.042

Reynolds, K.A., Mena, K.D., Gerba, C.P., 2008. Risk of waterborne illness via drinking water in the United States. Rev. Environ. Contam. Toxicol. 192, 117–158.

Ricci, A., Allende, A., Bolton, D., Chemaly, M., Davies, R., Fernandez Escamez, P.S., Herman, L., Koutsoumanis, K., Lindqvist, R., Nørrung, B., Robertson, L., Ru, G., Sanaa, M., Simmons, M., Skandamis, P., Snary, E., Speybroeck, N., Ter Kuile, B., Threlfall, J., Wahlström, H., Di Bartolo, I., Johne, R., Pavio, N., Rutjes, S., van der Poel, W., Vasickova, P., Hempen, M., Messens, W., Rizzi, V., Latronico, F., Girones, R., 2017. Public health risks associated with hepatitis E virus (HEV) as a food-borne pathogen. EFSA J. 15. https://doi.org/10.2903/j.efsa.2017.4886

Riveiro-Barciela, M., Mínguez, B., Gironés, R., Rodriguez-Frías, F., Quer, J., Buti, M., 2015. Phylogenetic demonstration of hepatitis E infection transmitted by pork meat ingestion. J. Clin. Gastroenterol. 49, 165–8. https://doi.org/10.1097/MCG.0000000000000113

Rivero-Juarez, A., Frias, M., Martinez-Peinado, A., Risalde, M.A., Rodriguez-Cano, D., Camacho, A., García-Bocanegra, I., Cuenca-Lopez, F., Gomez-Villamandos, J.C., Rivero, A., 2017. Familial

Hepatitis E Outbreak Linked to Wild Boar Meat Consumption. Zoonoses Public Health 1–5. https://doi.org/10.1111/zph.12343

Rosario, K., Duffy, S., Breitbart, M., 2012. A field guide to eukaryotic circular single-stranded DNA viruses: Insights gained from metagenomics. Arch. Virol. https://doi.org/10.1007/s00705-012-1391-y

Ross, T., 2008. Microbial ecology in food safety risk assessment, in: Schaffner, D. (Ed.), Microbiological Risk Assessment of Foods. ASM Press, Washington DC, pp. 51–98.

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. https://doi.org/10.1016/j.watres.2014.04.013

Rusiñol, M., Fernandez-Cassi, X., Timoneda, N., Carratala, 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. https://doi.org/10.1016/j.jenvman.2015.05.019

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. https://doi.org/10.1016/j.foodres.2014.08.018

Sano, D., Pérez-Sautu, U., Guix, S., Pintó, 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. https://doi.org/10.1128/AEM.01721-10

Santiago, P., Moreno, Y., Ferrús, M.A., 2015. Identification of Viable Helicobacter pylori in Drinking Water Supplies by Cultural and Molecular Techniques. Helicobacter 20, 252–259. https://doi.org/10.1111/hel.12205

Sanz, L.A., Gawlik, B.M., 2014. Water Reuse in Europe, Relevant guidelines, needs for and barriers to innovation Third, E-WAter (European Water Association). https://doi.org/10.2788/29234

Schijven, J., Rutjes, S., Smeets, P., Teunis, P., 2014. QMRAspot: a tool for quantitative microbial risk

assessment for drinking water Manual QMRAspot version 2.0, National Institute for Public Health and the Environment. The Netherlands.

Schijven, J.F., Teunis, P.F.M., Rutjes, S.A., Bouwknegt, M., de Roda Husman, A.M., 2011. QMRAspot: A tool for Quantitative Microbial Risk Assessment from surface water to potable water. Water Res. 45, 5564–5576. https://doi.org/10.1016/j.watres.2011.08.024

Shibata, S., Sekizuka, T., Kodaira, A., Kuroda, M., Haga, K., Doan, Y.H., Takai-todaka, R., Katayama, K., Wakita, T., Oka, T., Hirata, H., 2015. Complete Genome Sequence of a Novel GV.2 Sapovirus Strain, NGY-1, Detected from a Suspected Foodborne Gastroenteritis. Genome. Announc. 3, 2–3. https://doi.org/10.1128/genomeA.01553-14.Copyright

Shin, G.A., Sobsey, M.D., 2008. Inactivation of norovirus by chlorine disinfection of water. Water Res. 42, 4562–4568. https://doi.org/10.1016/j.watres.2008.08.001

Smeets, P.W.M.H., Rietveld, L.C., Van Dijk, J.C., Medema, G.J., 2010. Practical applications of quantitative microbial risk assessment (QMRA) for water safety plans. Water Sci. Technol. 61, 1561–1568. https://doi.org/10.2166/wst.2010.839

Souza, M., Azevedo, M.S.P., Jung, K., Cheetham, S., Saif, L.J., 2008. Pathogenesis and immune responses in gnotobiotic calves after infection with the genogroup II.4-HS66 strain of human norovirus. J. Virol. 82, 1777–1786. https://doi.org/10.1128/JVI.01347-07

Spandole, S., Cimponeriu, D., Berca, L.M., Mihaescu, G., 2015. Human anelloviruses: an update of molecular, epidemiological and clinical aspects. Arch. Virol. 160, 893–908. https://doi.org/10.1007/s00705-015-2363-9

Takanashi, S., Hashira, S., Matsunaga, T., Yoshida, A., Shiota, T., Tung, P.G., Khamrin, P., Okitsu, S., Mizuguchi, M., Igarashi, T., Ushijima, H., 2009. Detection, genetic characterization, and quantification of norovirus RNA from sera of children with gastroenteritis. J. Clin. Virol. 44, 161–163. https://doi.org/10.1016/j.jcv.2008.11.011

Teunis, P., Evers, E., Slob, W., 1999. Analysis of variable fractions resulting from microbial counts. Quant. Microbiol. 1, 63–88. https://doi.org/10.1023/A:1010028411716

Teunis, P., Moe, C.L., Liu, P., Miller, S., Lindesmith, L., Baric, B., Pendu, J., Calderon, R., 2008. Norwalk Virus: How infectious is it? Anticancer Res. 80, 1468–1476. https://doi.org/10.1002/jmv.21237

Teunis, P., Schijven, J., Rutjes, S., 2016. A generalized dose-response relationship for adenovirus infection and illness by exposure pathway. Epidemiol. Infect. 144, 3461–3473. https://doi.org/10.1017/S0950268816001862

Teunis, P.F.M., Havelaar, A.H., 2000. The Beta Poisson Dose-Response Model Is Not a Single-Hit Model 20.

Teunis, P.F.M., Rutjes, S.A., Westrell, T., de Roda Husman, A.M., 2009. Characterization of drinking water treatment for virus risk assessment. Water Res. 43, 395–404. https://doi.org/10.1016/j.watres.2008.10.049

Timoneda, N., 2017. Mètodes computacionals per a la identificaciò de virus emergents analizats per seqüenciaciò en massa en aigua y aliments. Universitat de Barcelona.

Todd, E., Greig, J., 2015. Viruses of foodborne origin: A review. Virus Adapt. Treat. 7, 25–45. https://doi.org/10.2147/VAAT.S50108

Touinssi, M., Gallian, P., Biagini, P., Attoui, H., Vialettes, B., Berland, Y., Tamalet, C., Dhiver, C., Ravaux, I., De Micco, P., De Lamballerie, X., 2001. TT virus infection: prevalence of elevated viraemia and arguments for the immune control of viral load. J. Clin. Virol. 21, 135–41.

Van Abel, N., Schoen, M.E., Kissel, J.C., Meschke, J.S., 2017. Comparison of Risk Predicted by Multiple Norovirus Dose-Response Models and Implications for Quantitative Microbial Risk Assessment. Risk Anal. 37, 245–264. https://doi.org/10.1111/risa.12616

Van der Kuyl, A.C., 2012. HIV infection and HERV expression: a review. Retrovirology 9, 6. https://doi.org/10.1186/1742-4690-9-6

van Duin, J., Olsthoorn, R.C.L., 2012. Family Leviviridae, in: King, A., Adams, M., Carstens, E., Lefkowitz, E. (Eds.), Virus Taxonomy, Clasification and Nomenclature of Viruses. ElSevier Academic Press, Davis, pp. 1035–1043.

Vose, D., 2008. Risk analysis a quantitative guide, third edit. ed. John Wiley and Sons, Wiltshire, England.

WHO, 2016a. Quantitative Microbial Risk Assessment: Application for Water Safety Management. Geneva, Switzerland. https://doi.org/10.1002/9781118910030

WHO, 2016b. Technical considerations and case definitions to improve surveillance for viral

hepatitis: technical report. Geneve.

WHO, 2014. Preventing diarrhoea through better water, sanitation and hygiene, WHO. Geneva, Switzerland.

WHO, 2006. WHO guidelines for the safe use of wastewater , excreta and greywater: Wastewater use in agriculture II, 204. https://doi.org/10.1007/s13398-014-0173-7.2

Wold, W.S.., Horwitz, M.., 2013. Adenoviruses, in: Knipe, D.M., Howley, P.M. (Eds.), Fields Virology. Lippincott Williams & Wilkins, Philadelphia, pp. 1732–1767.

World Health Organization, 2016. Global health sector strategy on viral hepatitis 2016–2021, Towards ending viral hepatitis. Geneva, Switzerland.

Wyn-Jones, A.P., Carducci, A., Cook, N., D'Agostino, M., Divizia, M., Fleischer, J., Gantzer, C., Gawler, A., Girones, R., Höller, C., de Roda Husman, A.M., Kay, D., Kozyra, I., López-Pila, J., Muscillo, M., José Nascimento, M.S., Papageorgiou, G., Rutjes, S., Sellwood, J., Szewzyk, R., Wyer, M., 2011. Surveillance of adenoviruses and noroviruses in European recreational waters. Water Res. 45, 1025–1038. https://doi.org/10.1016/j.watres.2010.10.015

Xiao, J., Li, J., Hu, G., Chen, Z., Wu, Y., Chen, Y., Chen, Z., Liao, Y., Zhou, J., Ke, X., Ma, L., Liu, S., Zhou, J., Dai, Y., Chen, H., Yu, S., Chen, Q., 2011. Isolation and phylogenetic characterization of bat astroviruses in southern China. Arch. Virol. 156, 1415–1423. https://doi.org/10.1007/s00705-011-1011-2

Yokoyama, C.C., Loh, J., Zhao, G., Stappenbeck, T.S., Wang, D., Huang, H. V., Virgin, H.W., Thackray, L.B., 2012. Adaptive Immunity Restricts Replication of Novel Murine Astroviruses. J. Virol. 86, 12262–12270. https://doi.org/10.1128/JVI.02018-12

Zhang, Y., Li, F., Shan, T.-L., Deng, X., Delwart, E., Feng, X.-P., 2016. A novel species of torque teno mini virus (TTMV) in gingival tissue from chronic periodontitis patients. Sci. Rep. 6, 26739. https://doi.org/10.1038/srep26739

Zur Hausen, H., Bund, T., de Villiers, E.-M., 2017. Infectious Agents in Bovine Red Meat and Milk and Their Potential Role in Cancer and Other Chronic Diseases. Curr. Top. Microbiol. Immunol. 1–34. https://doi.org/10.1007/82_2017_3

8 ANEXOS

8.1 Supplementary material paper: Characterization of the efficiency and uncertainty of skimmed milk flocculation for the simultaneous concentration and quantification of water-borne viruses, bacteria and protozoa.

```
#------------------------------------------------------------ clean all objects & load package boot
rm(list=ls(all=TRUE))
gc();
library(boot)

#---------------------------- initial seed to replicate paper's results. Comment for different runs
set.seed(3082016)

m1BootPred <- function(data,indices) {
  d <- data[indices]
  Y <- sample(data,size=1)
  direct <- (mean(d)-Y) / sd(d)
  data2  <- 1/data
  d <- data2[indices]
  Y <- sample(data2,size=1)
  inverse <- (mean(d)-Y) / sd(d)
  return(c(direct,inverse))
}

computeBoot_T_plusNormal <- function(actualData, conf.level=0.95) {

  alpha <- 1-conf.level
  qmax  <- 1-(alpha/2)
  qmin  <- alpha/2

  # Boot-t pred intervals
  nS <- length(actualData)
  resultsBootT <- boot(data=actualData, statistic=m1BootPred,R=1000)
  ruMin <- mean(actualData) - sd(actualData)* quantile(resultsBootT$t[,1],probs=c(qmax)) # *
sqrt(1+1/nS)
  ruMax <- mean(actualData) - sd(actualData)* quantile(resultsBootT$t[,1],probs=c(qmin)) # *
sqrt(1+1/nS)
  m5 <- c(ruMin,ruMax)

  # standard normal prediction values
  normMin <- mean(actualData) - sd(actualData)* sqrt(1+1/nS) * qt(p=c(qmax),df=(nS-1))
  normMax <- mean(actualData) + sd(actualData)* sqrt(1+1/nS) * qt(p=c(qmax),df=(nS-1))
  m6 <- c(normMin,normMax)

  m7<- c(1/ruMax,1/ruMin)        # concentration reverses the PI bounds of the recovery
  m8<- c(1/normMax,1/normMin)

  m9  <- m5 / mean(actualData)   # relative error ratio
  m10 <- m6 / mean(actualData)


  results <- data.frame(m5,m6,m7,m8,m9,m10)
  colnames(results) <- c("Boot-t (recov.)", "Normal (recov.)", "Boot-t (conc.)", "Normal
(conc.)","Boot-t (rel.)", "Normal (rel.)")

  rownames(results) <- c(paste(conf.level,"% PI lower bound", sep=''),paste(conf.level,"% PI upper
bound", sep=''))

  return(results)
}

Rotavirus    <- c(30,34,32,28,37,31,35,33,25,23,29,30,29,22,16,31,24,23,25)/100
computeBoot_T_plusNormal (Rotavirus,0.95)
```

8.2 Supplementary materials: QMRA manuscript

Table S1: raw database

WWTP	month	Raw sewage		Secondary		Tertiary	
		HAdV	NoVII	HAdV	NoVII	HAdV	NoVII
1	1	1164160	28262200	126030	3124250	322,47	21566
1	2	162933,33	20667	18033,33	11400	423,64	69,666667
1	3	135550	227675	29906,67	0	191,23	0
1	4	105933,33	1794666,7	41050	113600	965	0
1	5	50100	1850000	4070	0	311	0
1	6	60400	370905	9585	64500	37,9	18,85
1	7	281000	9815000	4550	191850	18,1	155,5
1	8	234000	16850000	6120	99050	0	0
1	9	22110	335000	5280	4615	1115	25200
1	10	12660	993500	2870	148500	1775	50300
1	11	68866,67	419400	0	26466,667	0	31350
1	12	78356,67	1061000	0	18450	0	69050
2	4	1,85E+04	1,25E+05	3,67E+03	0	5,75E+01	0
2	5	1,84E+04	1,76E+04	6,06E+01	3,67E+02	0	0
2	6	1,23E+04	2,33E+06	7,60E+02	2,93E+05	1,32E+01	0
2	7	1,73E+04	1,48E+03	1,38E+02	2,54E+03	0	4,05E+02
2	8	1,58E+05	1,86E+05	9,52E+03	6,77E+04	0	0
2	10	2,64E+04	2,38E+04	5,24E+02	0	0	0
2	11	2,98E+04	9,67E+06	6,69E+02	1,01E+05	0	0
2	12	4,80E+04	5,38E+06	6,52E+02	4,20E+04	0	6,14E+01
2	1	1,77E+05	7,77E+04	8,99E+04	6,92E+04	8,24E+02	5,20E+02
2	2	5,64E+04	2,17E+05	4,50E+03	0	2,90E+01	0
2	3	2,07E+05	9,14E+06	1,56E+03	3,48E+03	0	0
2	4	3,68E+04	3,92E+05	3,43E+03	1,79E+05	0	0

Table S2: Log reduction of WWTP efficiency by each virus with mean concentration (percentile 95) in actual situation and log reduction values with concentration in tertiary effluent required to reach suggestions of WHO (10^{-6} DALYs).

		WWTP1		WWTP2	
		HAdV	NoV GII	HAdV	NoV GII
Actual	Mean	3.7	5.2	8.9	12.2
	5%	1.8	1.1	2.1	3.2
	50%	3.2	4	6.8	9.3
	95%	7.9	13.6	23.2	30.8
To reach 10^{-6} DALYs	Mean	6.7	10	11.5	14.6
	5%	4.8	6.1	4.5	5.7
	50%	6.2	9	9.3	11.8
	95%	10.5	18.3	25.6	33

Table S3: FAST analysis as sensitivity analysis for the aggregated and non-aggregated model (Conservative scenario) for NOV GII and with $P_{ill|inf}$ estimated with eq 5. Values represent the total order estimation between input parameters and daily probability of illness (Pill).

Parameters	WWTP 1		WWTP2	
	HAdV	NoV	HAdV	NoV
g	0.834	0.799	0.824	0.869
π	0.872	0.874	0.875	0.875
R_s	0.567	0.557	0.631	0.868
R_t	0.567	0.557	0.631	0.868
R_{wash}	0.755	0.652	0.688	0.875
V_{surf}	0.564	0.480	0.642	0.763
R_{inf}	0.567		0.631	
π_{rec}	0.324	0.355	0.564	0.811
Ing	0.844	0.811	0.855	0.847

Table S4: FAST analysis as sensitivity analysis for the non-aggregated model (Worst-case scenario) for NOV GII and with $P_{ill|inf}$ estimated with eq 5. Values represent the total order estimation between input parameters and daily probability of illness (Pill).

Parameters	WWTP 1		WWTP2	
	HAdV	NoV	HAdV	NoV
g	0.566	0.320	0.551	0.609
Π	0.794	0.767	0.875	0.878
R_s	0.282	0.059	0.301	0.278
R_t	0.282	0.059	0.301	0.278
R_{wash}	0.408	0.085	0.375	0.346
V_{surf}	0.207	0.019	0.290	0.188
R_{inf}	0.282		0.301	
π_{rec}	0.067	0.006	0.177	0.174
Ing	0.703	0.356	0.722	0.624

Table S5: Spearman correlation as sensitivity analysis for the aggregated and non-aggregated model (Conservative scenario) for NOV GII and with $P_{ill|inf}$ estimated with eq 5. Values represent a Spearman correlation between input parameters and daily probability of illness (Pill).

Parameters	WWTP 1		WWTP2	
	HAdV	NoV	HAdV	NoV
g	0.2745	0.2608	0.0933	0.1881
π	0.7139	0.8769	0.9515	0.9233
C_{eff}	0.7765	0.9333	0.95786	0.9533
V_{surf}	0.06487	0.0364	0.0264	0.01928
R_s	-0.14025	-0.0813	-0.0586	-0.0433
R_t	-0.14257	-0.0824	-0.0566	-0.04011
R_{wash}	-0.1793	-0.1019	-0.0721	-0.0551
R_{inf}	-0.1408	-	-0.0542	-
π_{rec}	-0.013745	-0.0052	0.00586	-0.0046
Ing	0.48285	0.2742	0.19659	0.1467

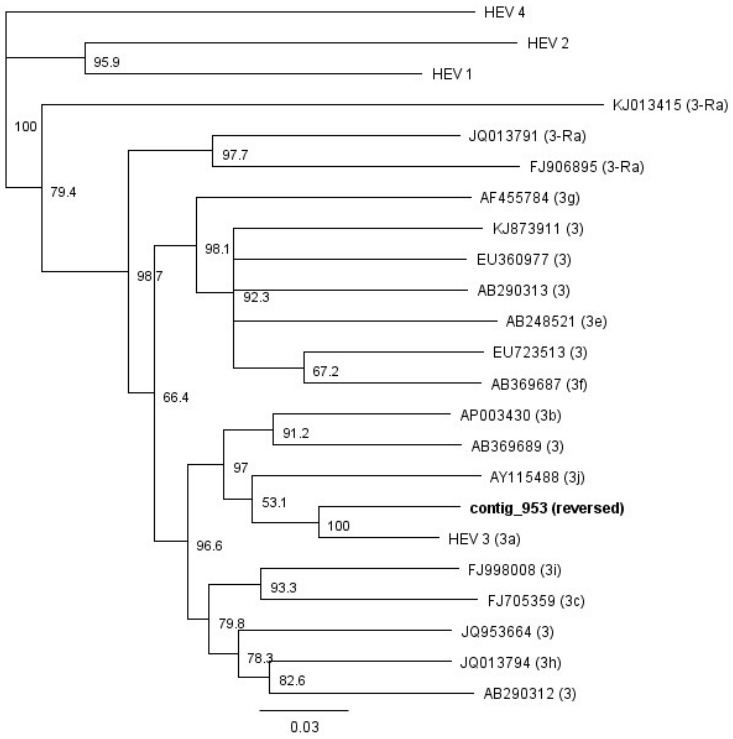
Table S6: Spearman correlation as sensitivity analysis for the non-aggregated model (worst-case scenario) for NOV GII and with $P_{ill|inf}$ estimated with fix value 0.5 and 0.7 for HAdV and NoV GII, respectively. Values represent a Spearman correlation between input parameters and daily probability of illness (Pill).

Parameters	WWTP 1		WWTP2	
	HAdV	NoV	HAdV	NoV
g	0.2732	0.2635	0.0905	0.1823
π	0.7111	0.8762	0.9539	0.9413
C_{eff}	0.7784	0.9329	0.9600	0.9712
V_{surf}	0.0649	0.0395	0.0280	0.0158
R_s	-0.1420	-0.0823	-0.0577	-0.0441
R_t	-0.1407	-0.0823	-0.0590	-0.0420
R_{wash}	-0.1746	-0.0989	-0.0690	-0.0539
R_{inf}	-0.1416		-0.0537	
π_{rec}	-0.0122	-0.0074	-0.0064	-0.0075
Ing	0.4796	0.2705	0.1934	0.1444

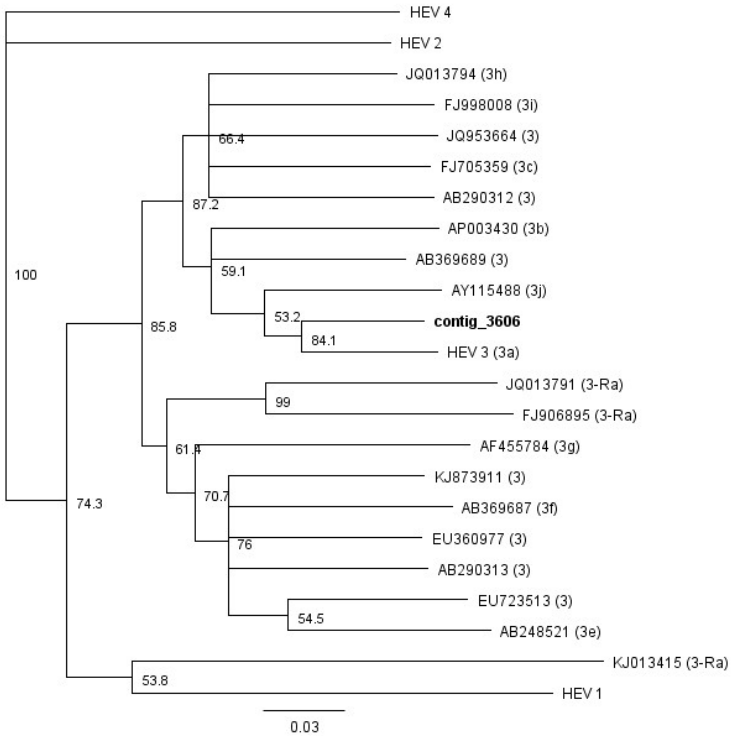
8.3 S1 Supporting Information: Individual phylogenetic trees computed from contigs over reference genome locations in HEV.

HEV SH5

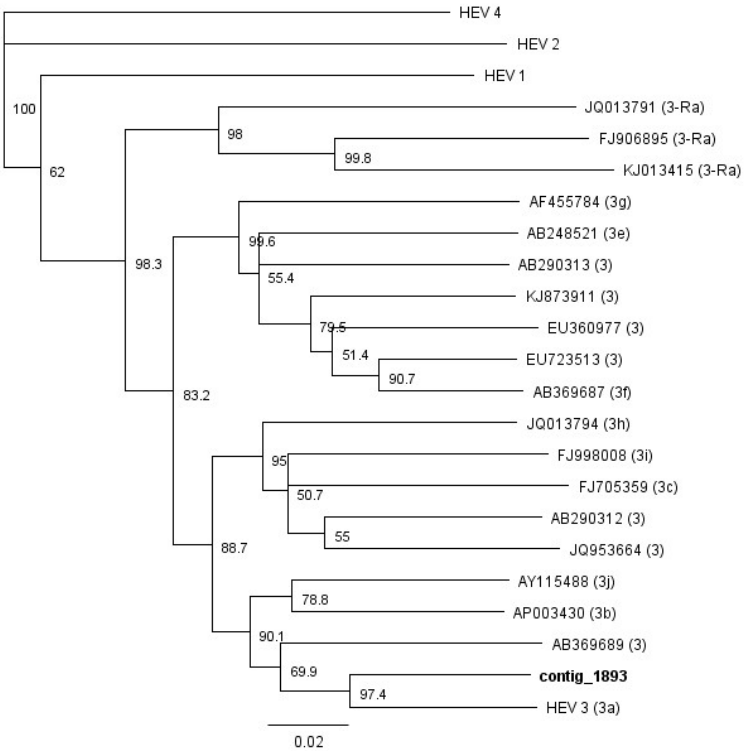
Contig_953



Contig 3606

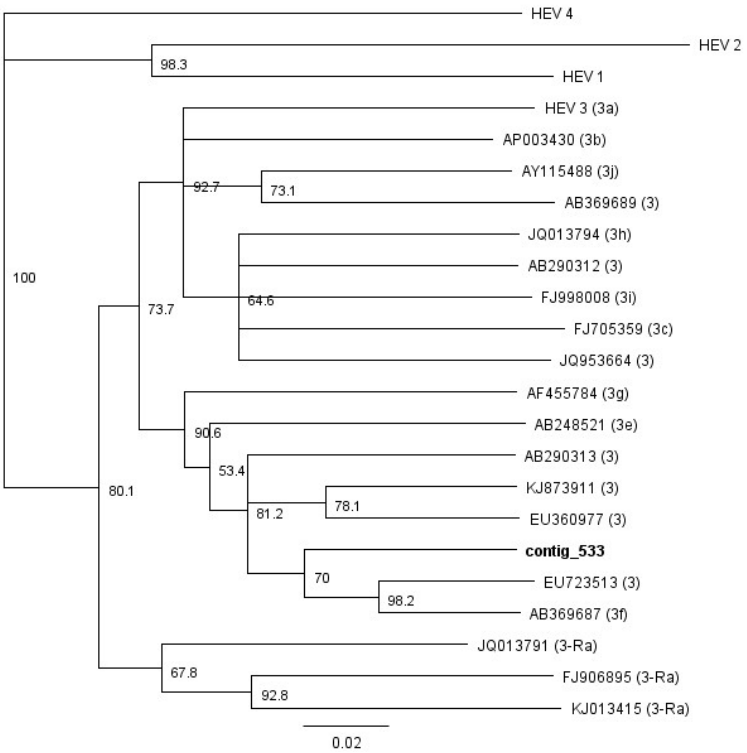


Contig 1893

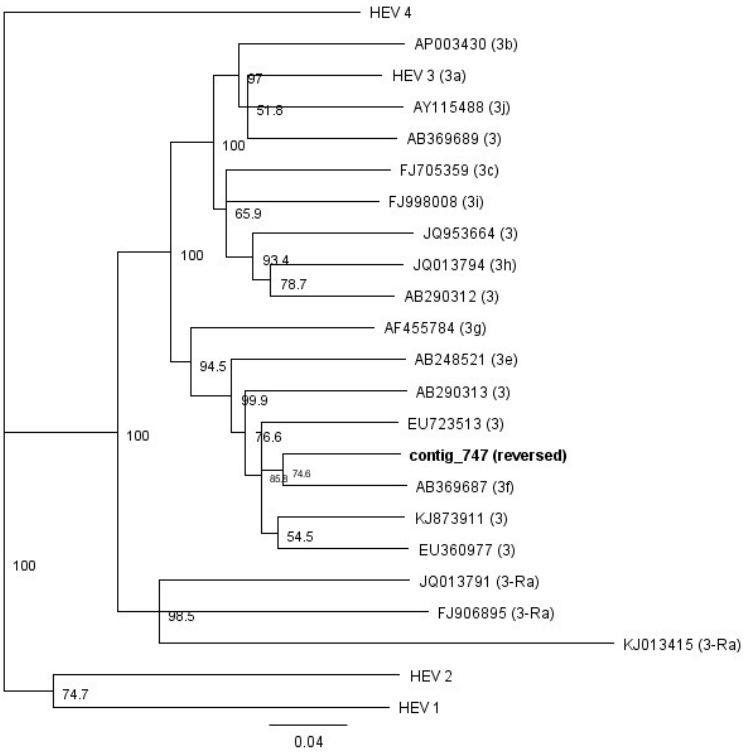


HEV SH6

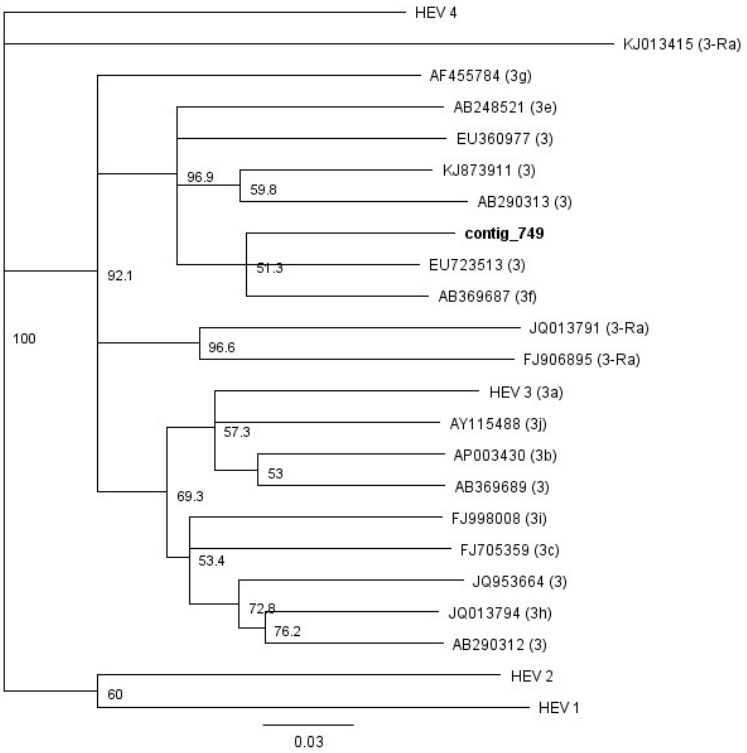
Contig 533



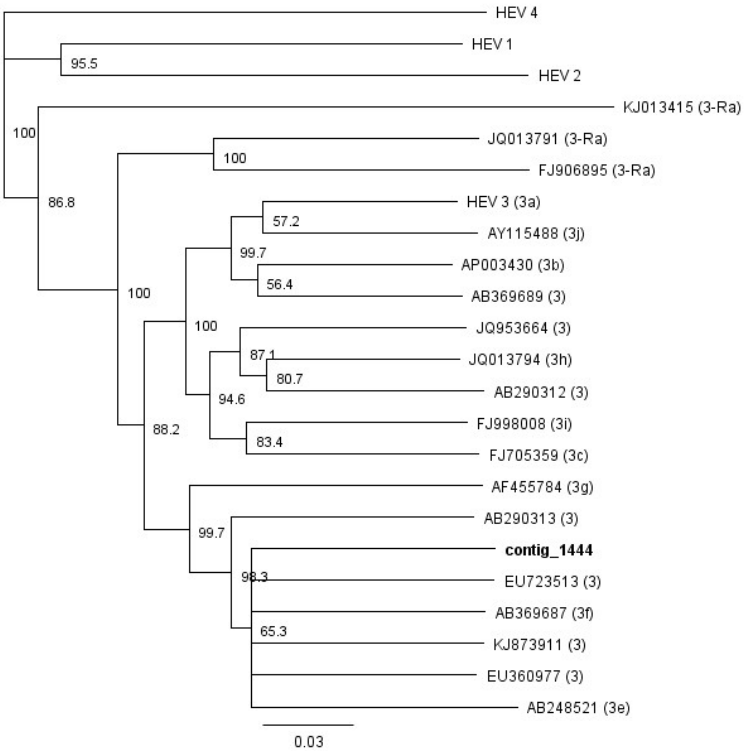
Contig 747



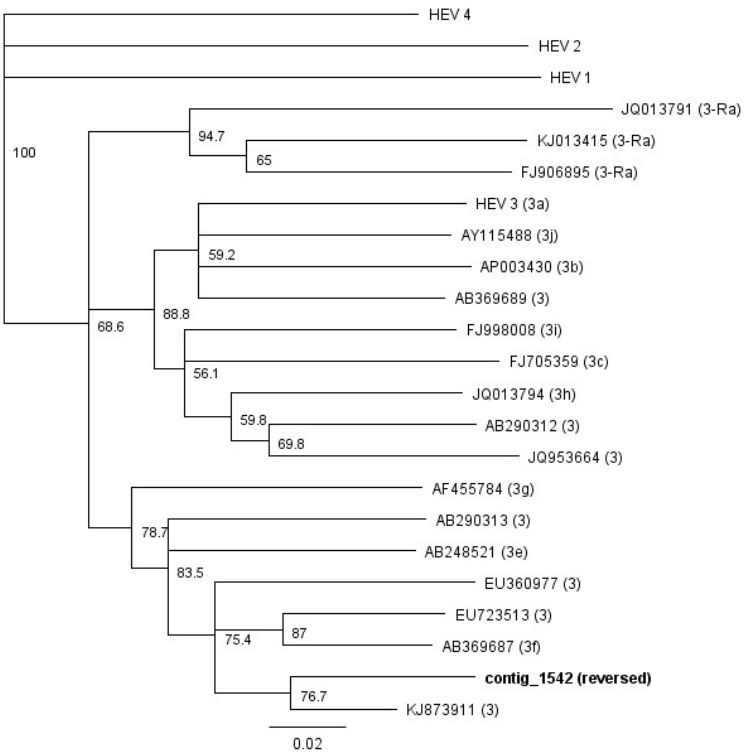
Contig 749



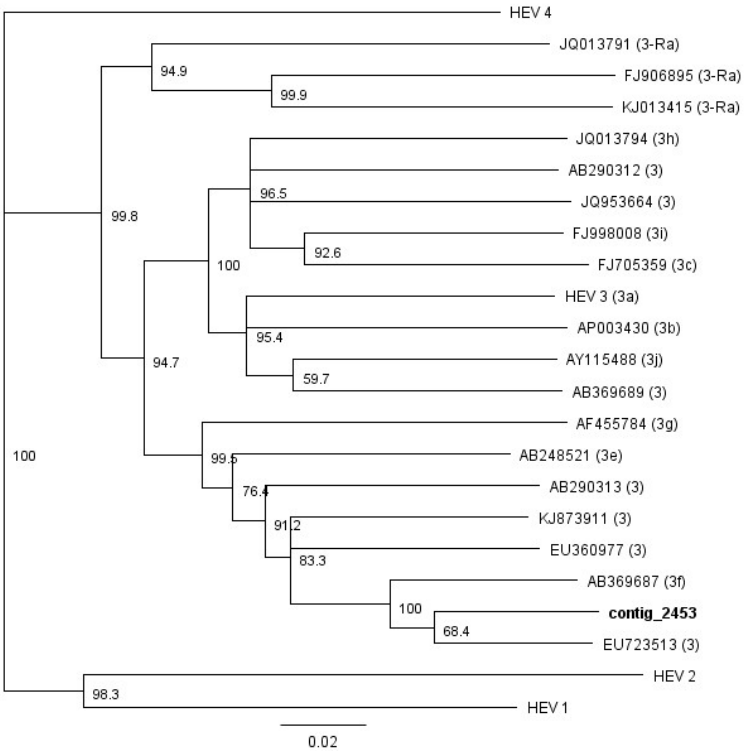
Contig 1444



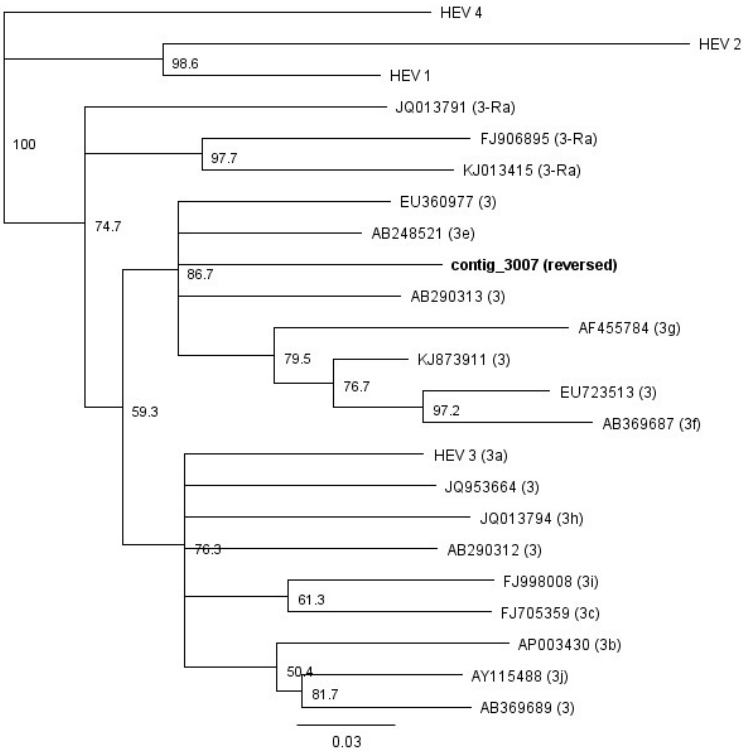
Contig 1542



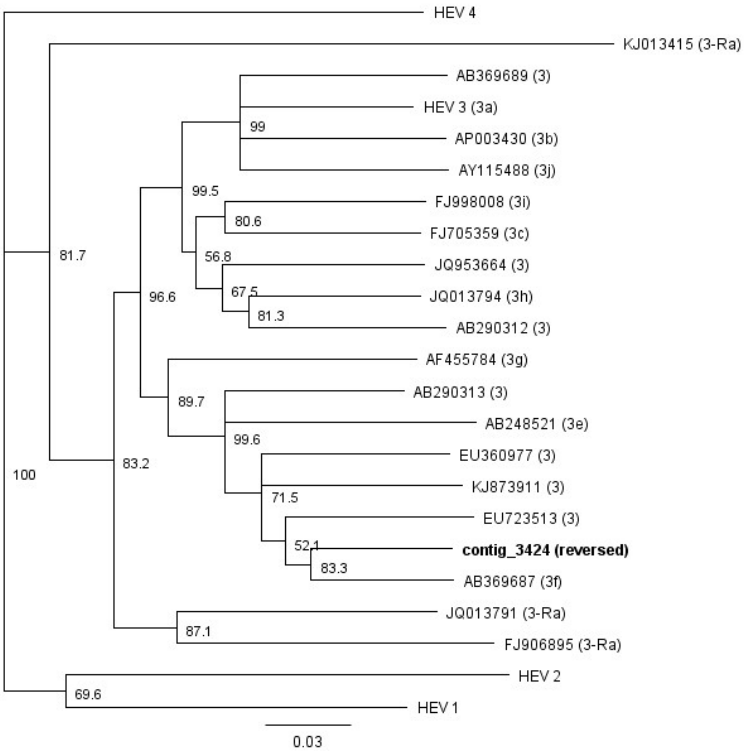
Contig 2453



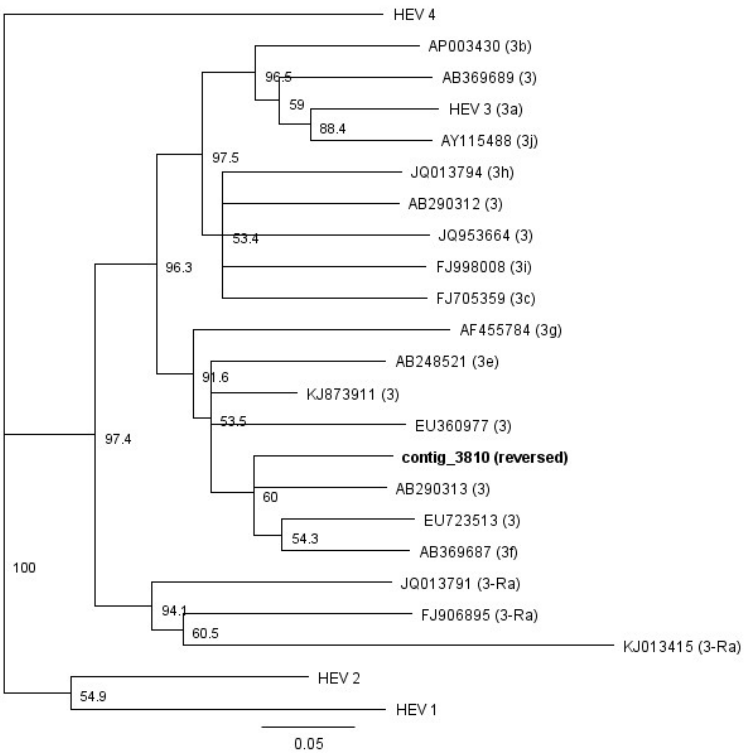
Contig 3007



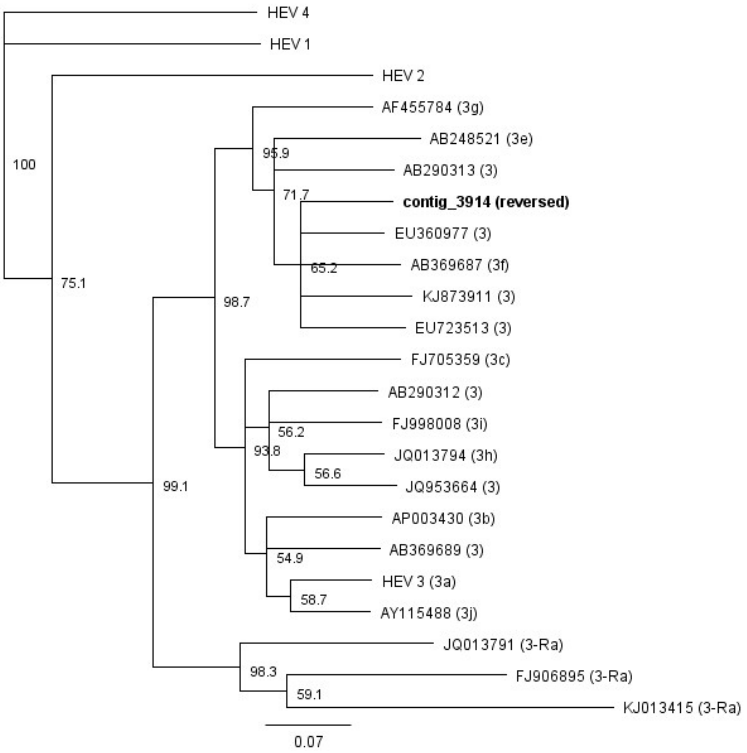
Contig 3424



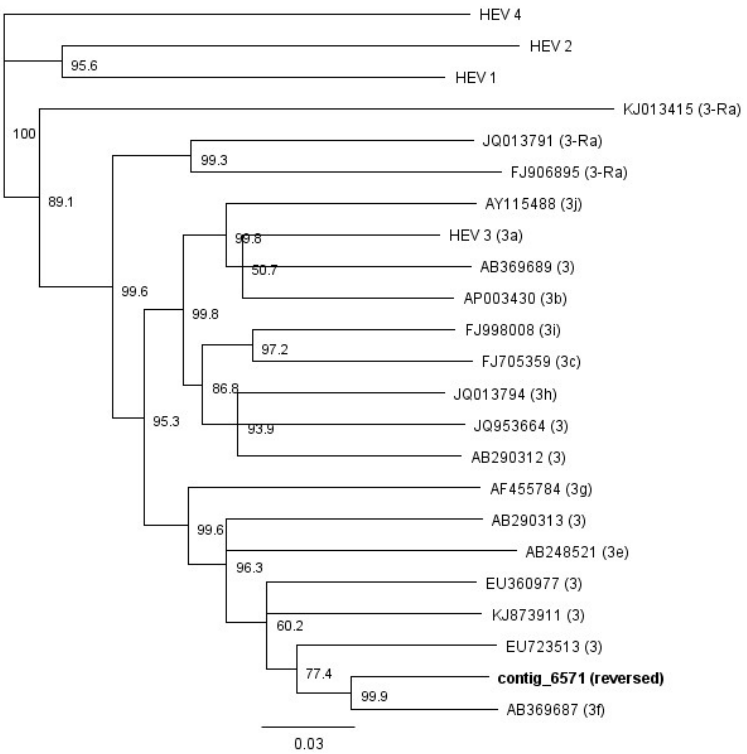
Contig 3810



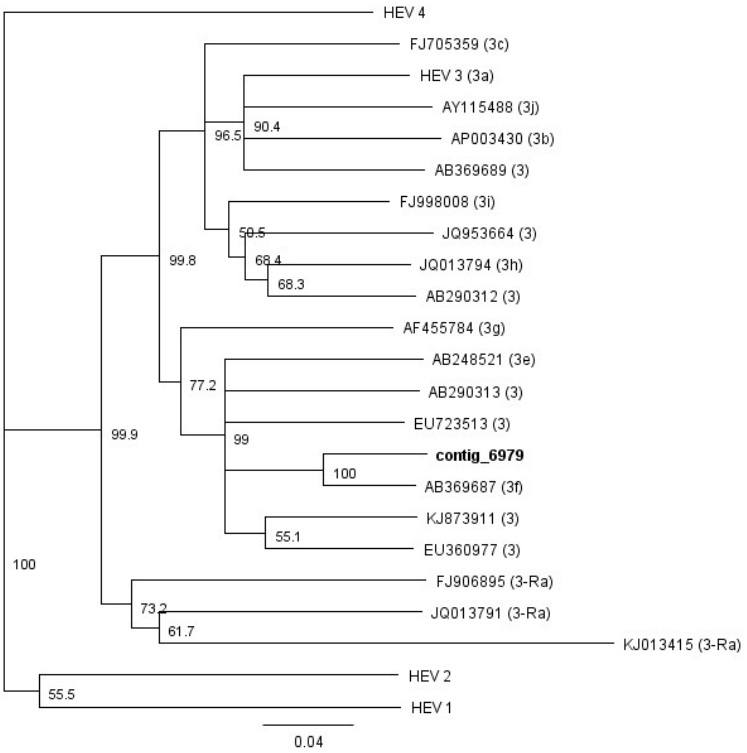
Contig 3914



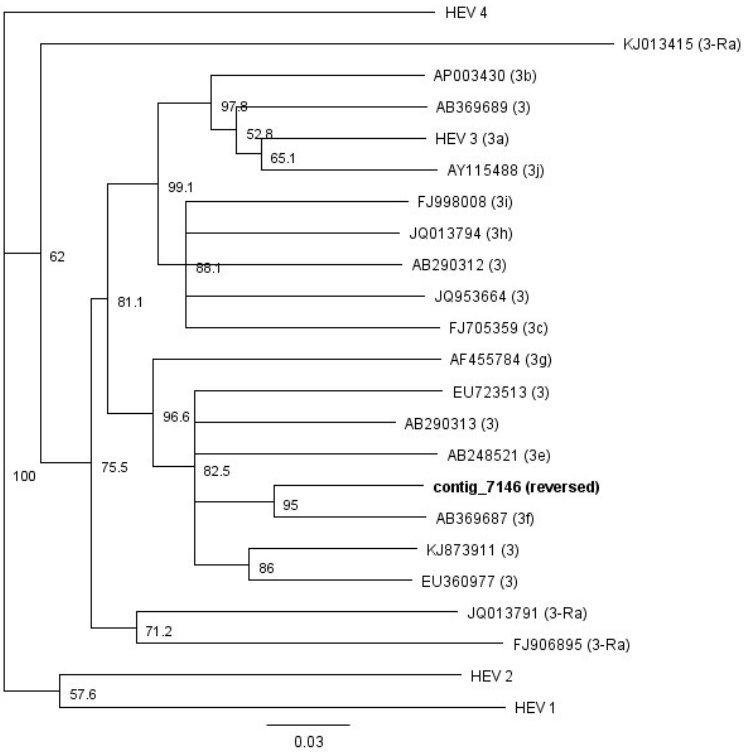
Contig 6571



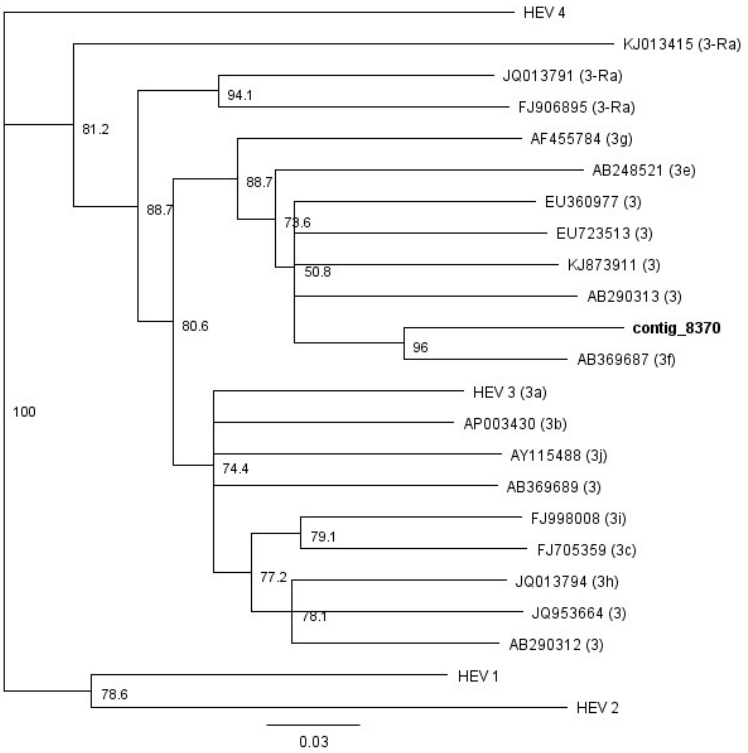
Contig 6979



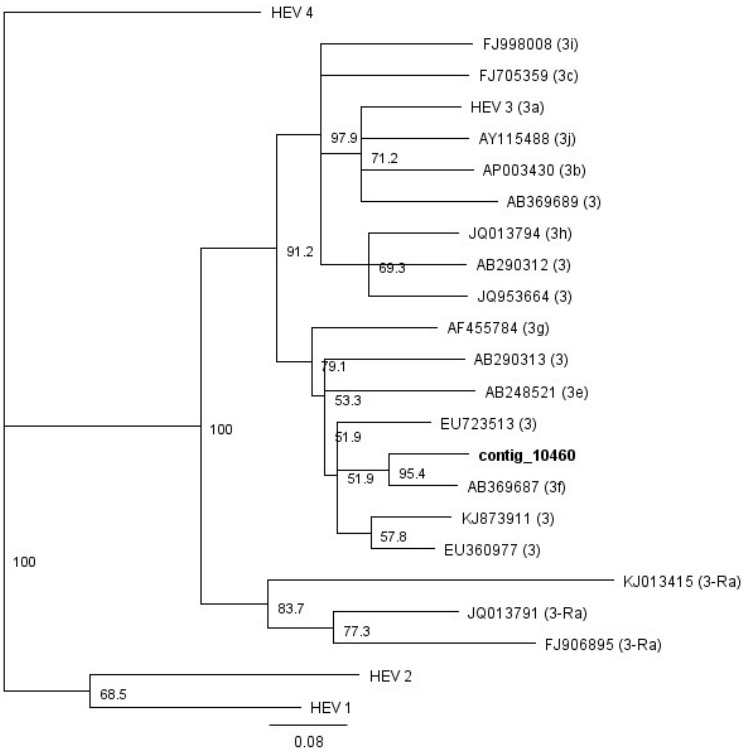
Contig 7146



Contig 8370



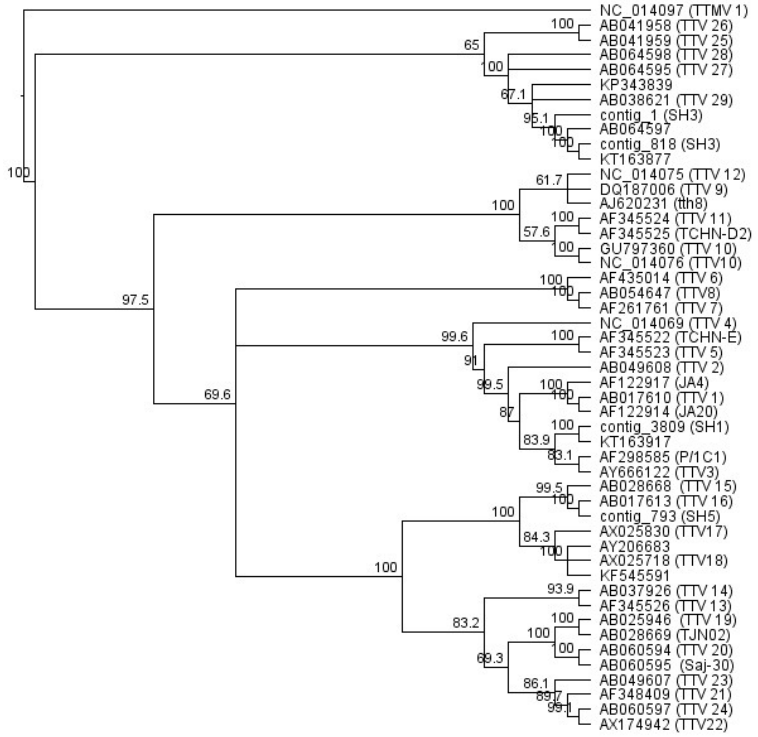
Contig 10460



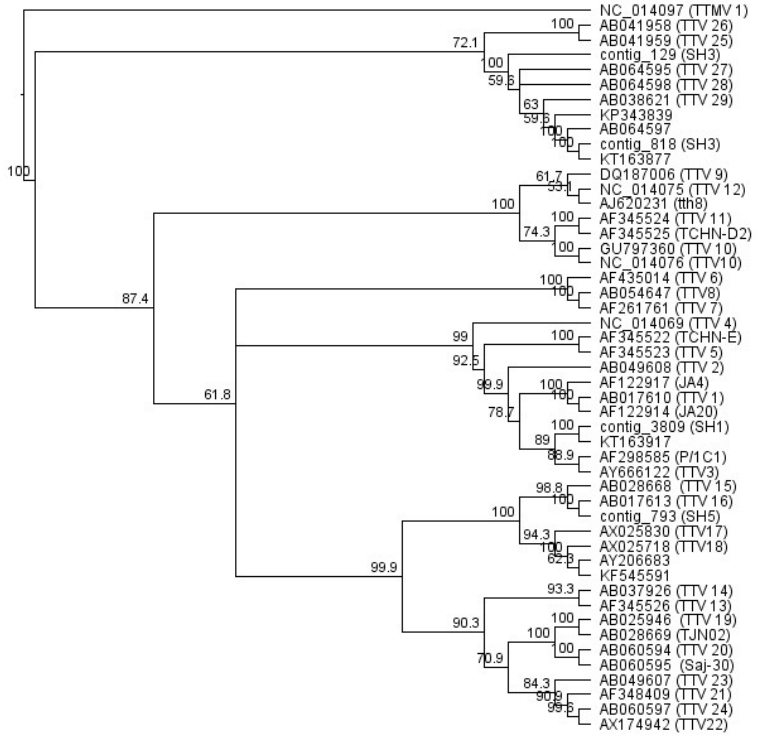
8.4 S2 Supporting Information: Individual phylogenetic trees computed from contigs over reference genome locations in *Anelloviridae* family.

TTV

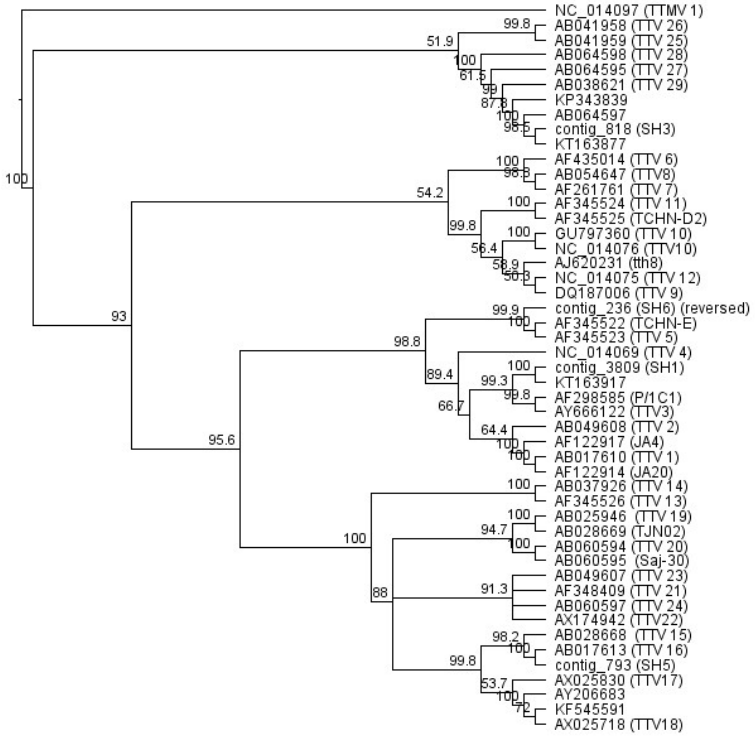
Contig 1



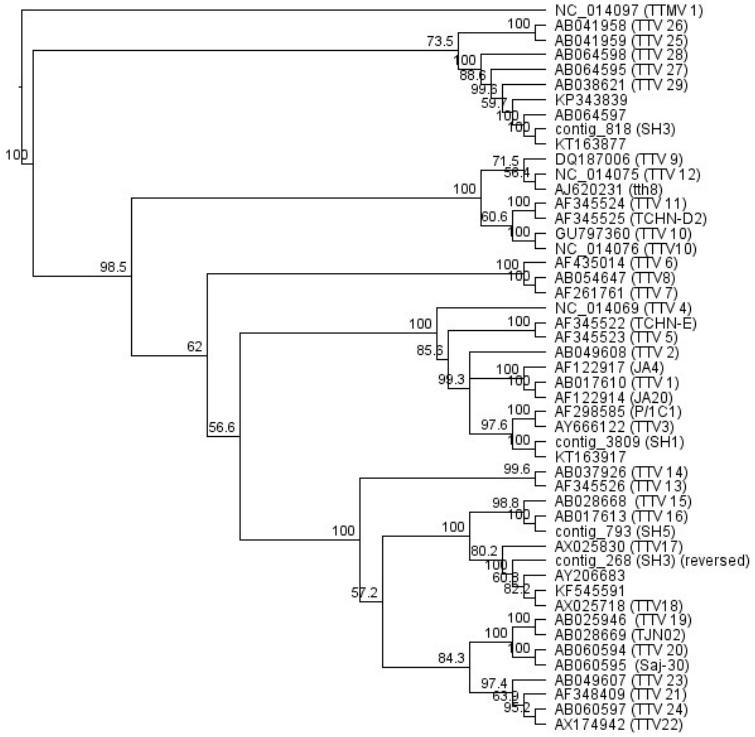
Contig 129



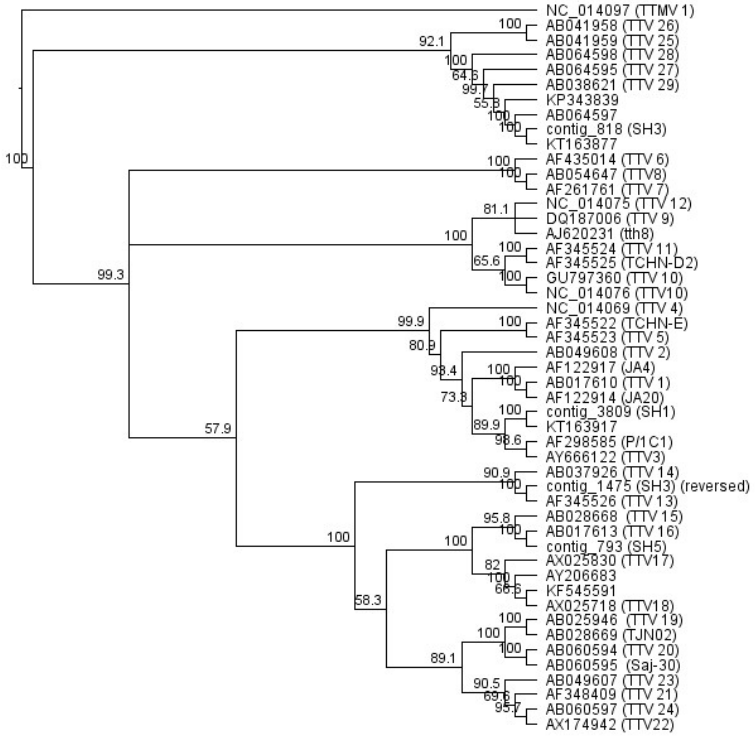
Contig 236



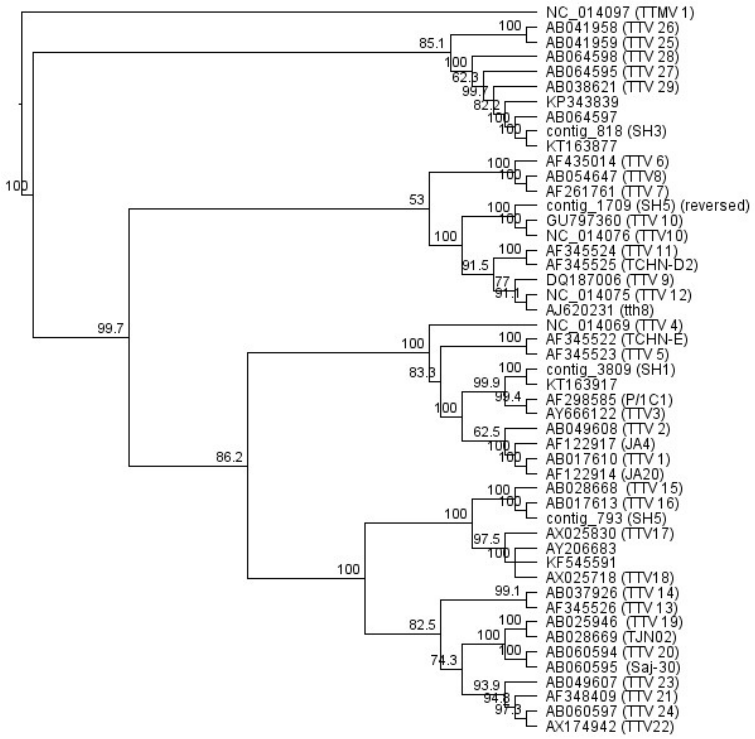
Contig 268



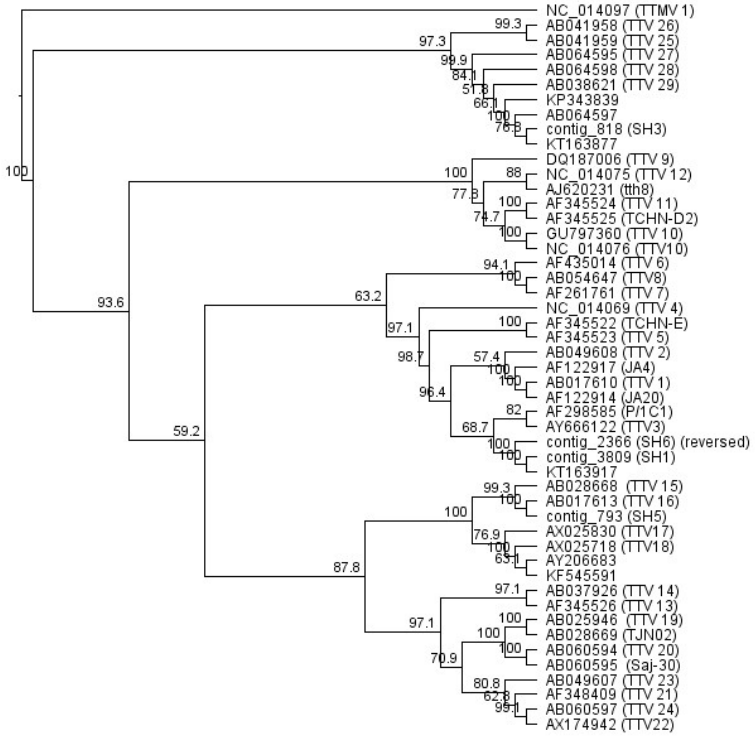
Contig 1475



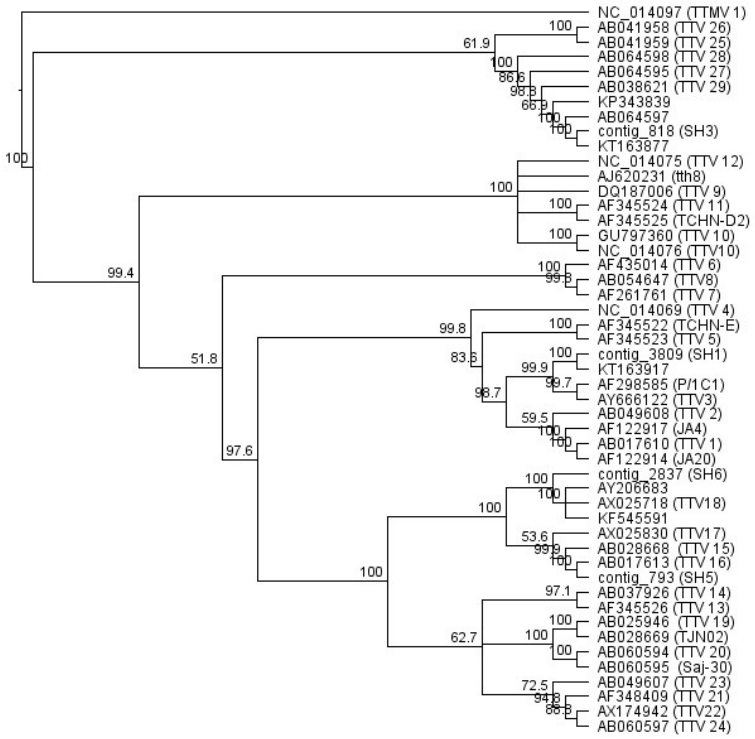
Contig 1709



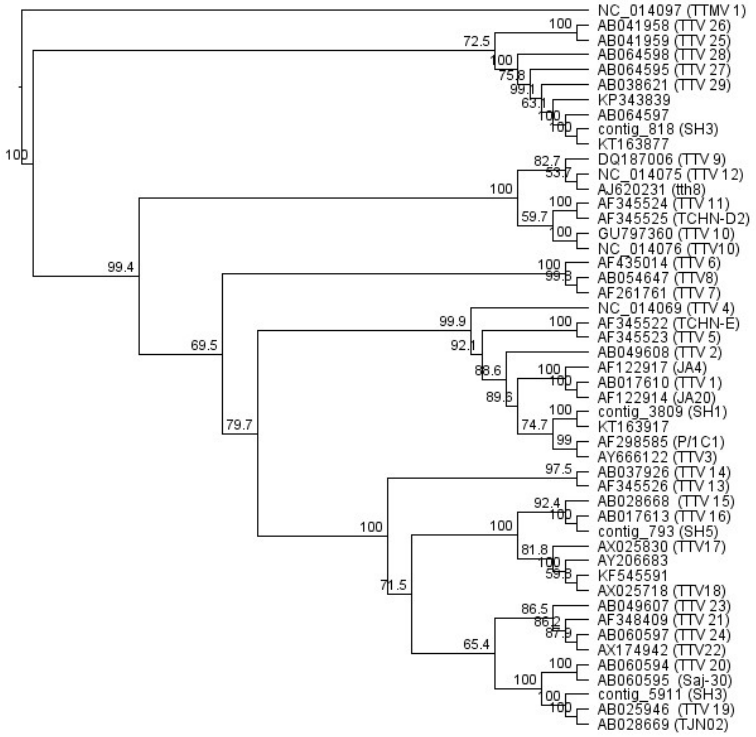
Contig 2366



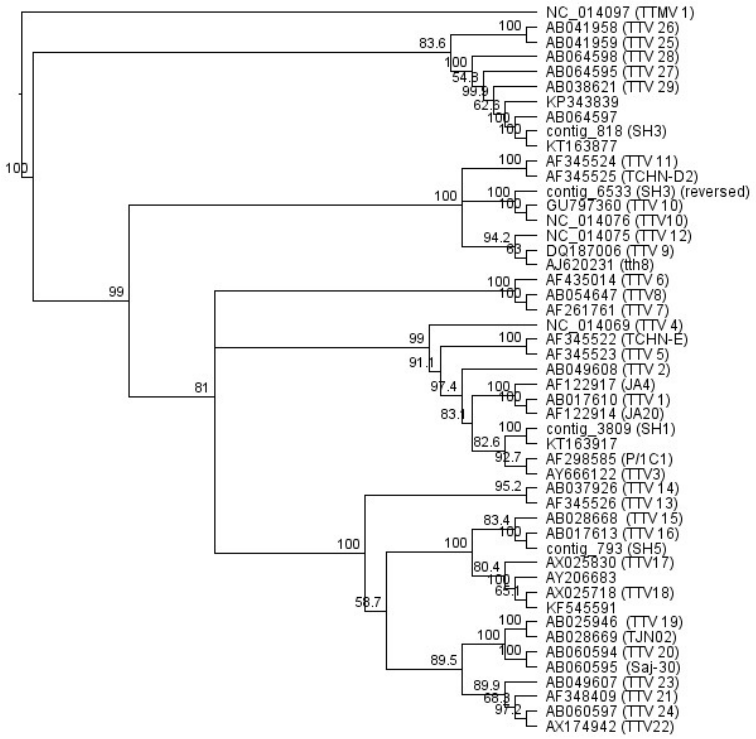
Contig 2837



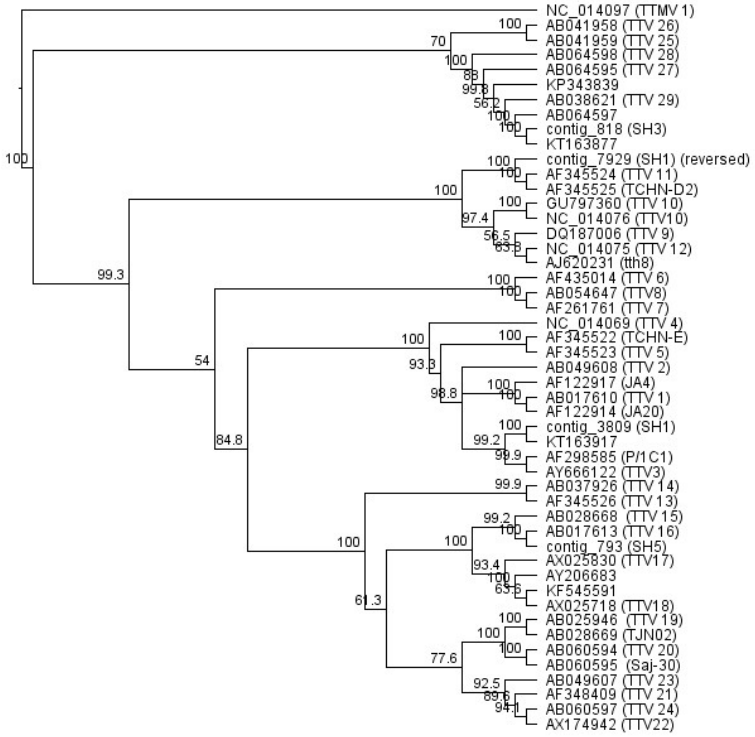
Contig 5911



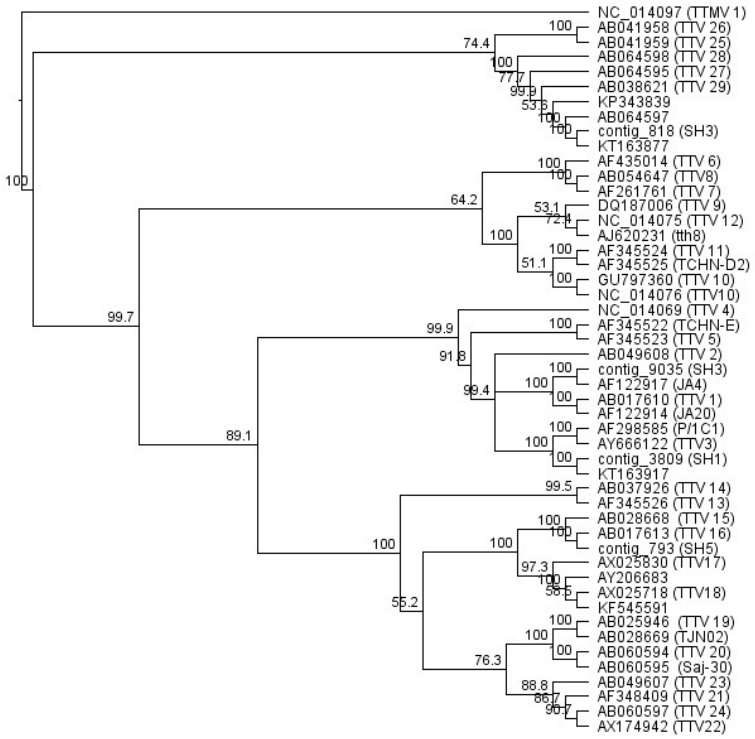
Contig 6533



Conti 7929

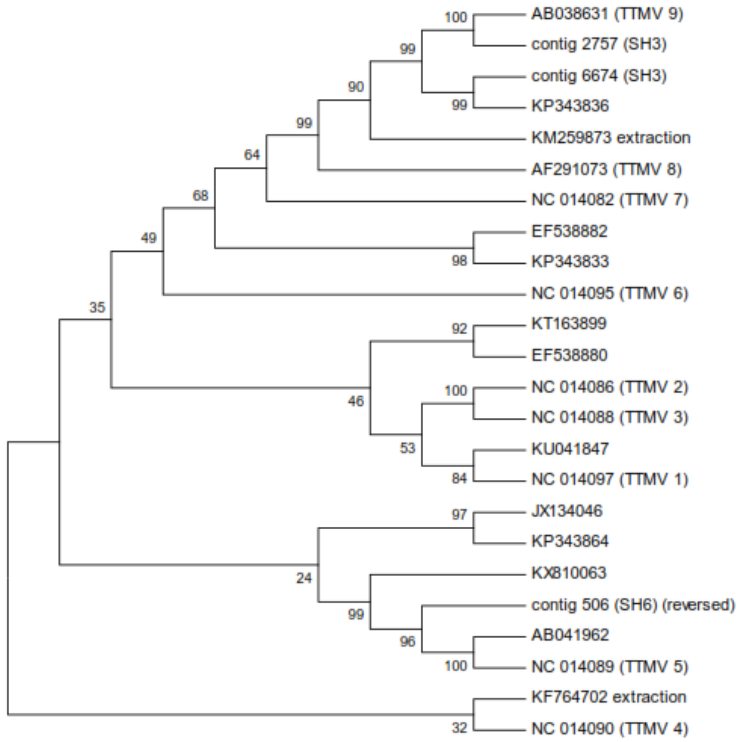


Contig 9035

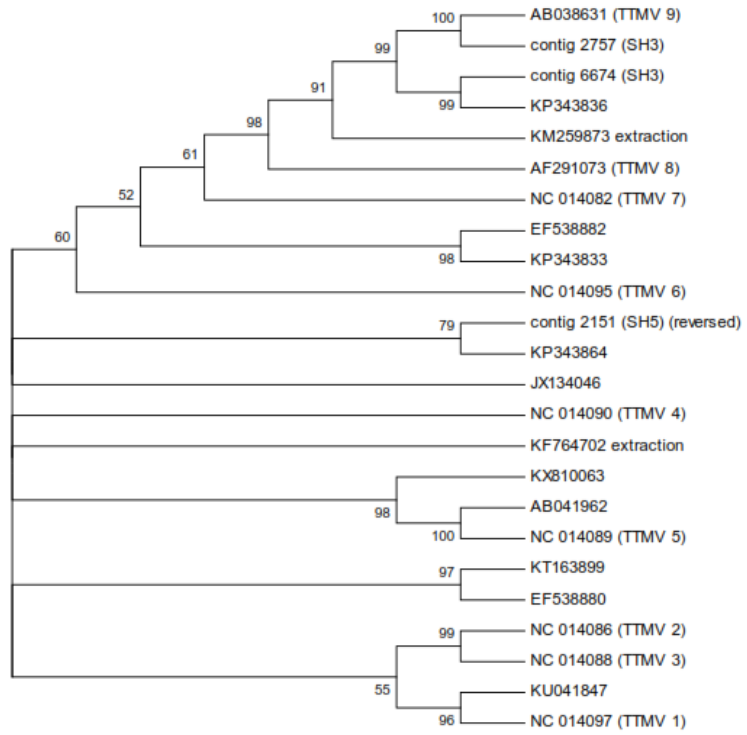


TTMV

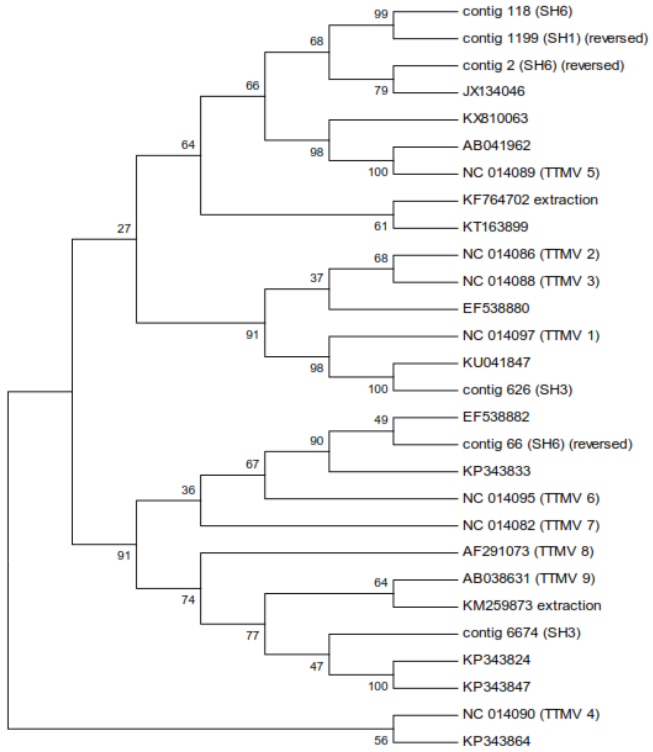
Contig 506



Contig 2151

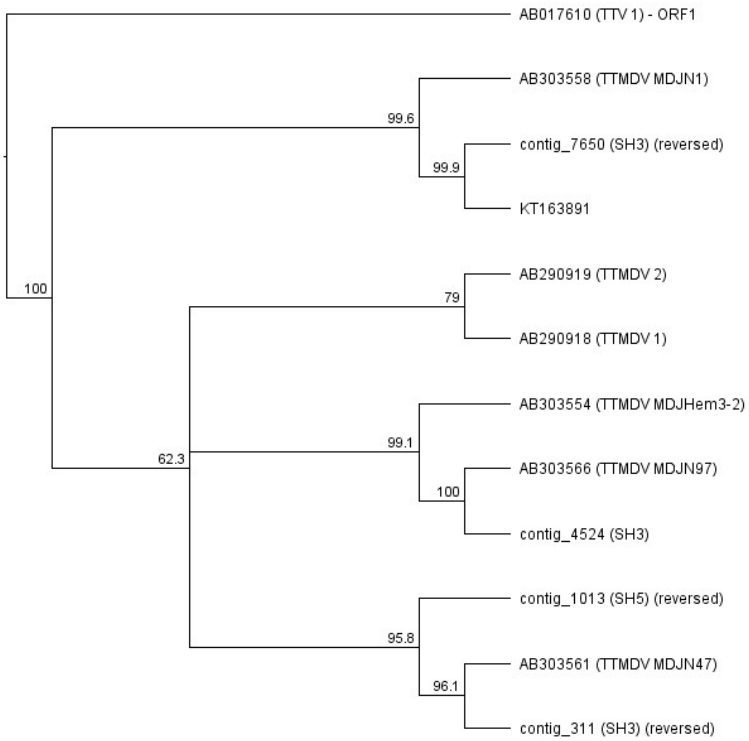


TTMV Extended región (Letters in the original phylogenetic tree)

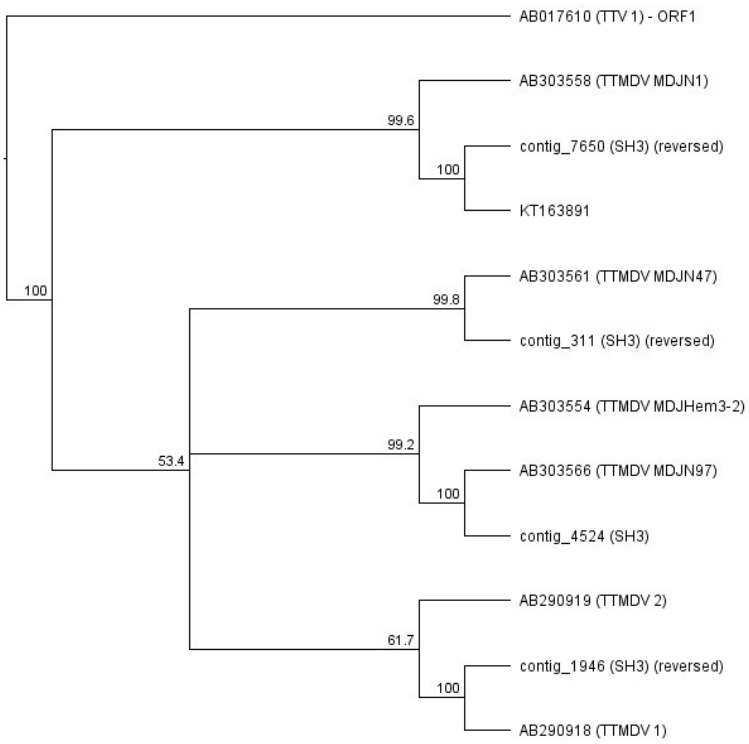


TTMDV

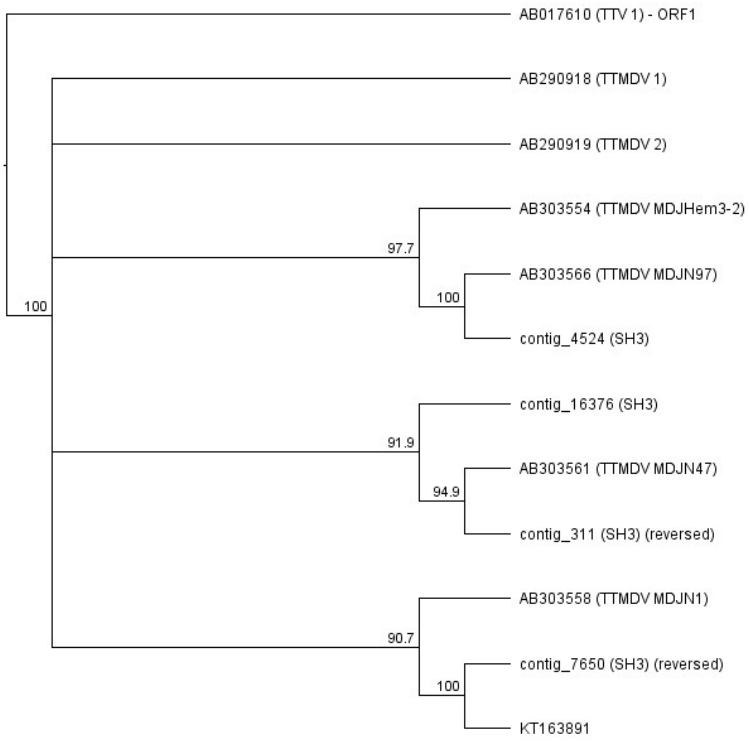
Contig 1013



Contig 1946

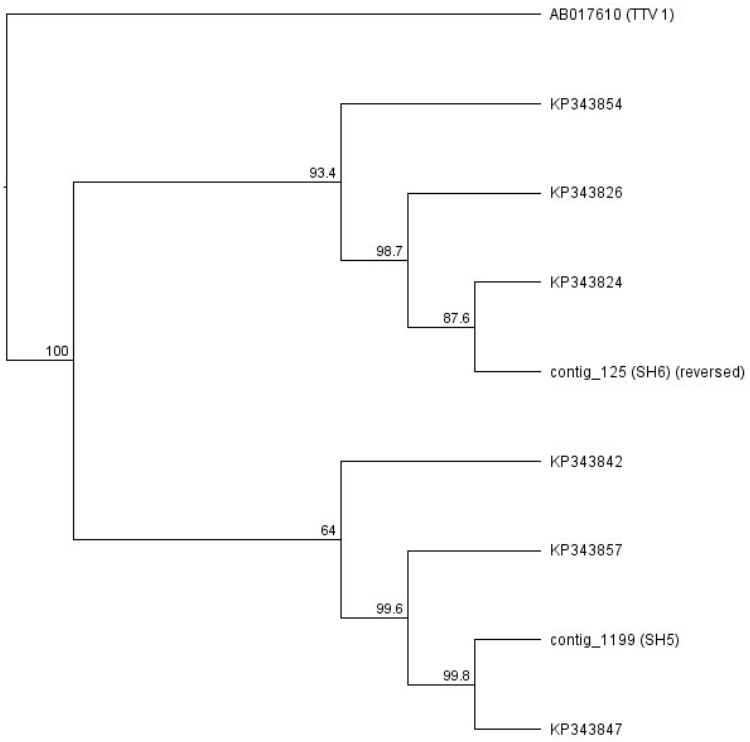


Contig 16376

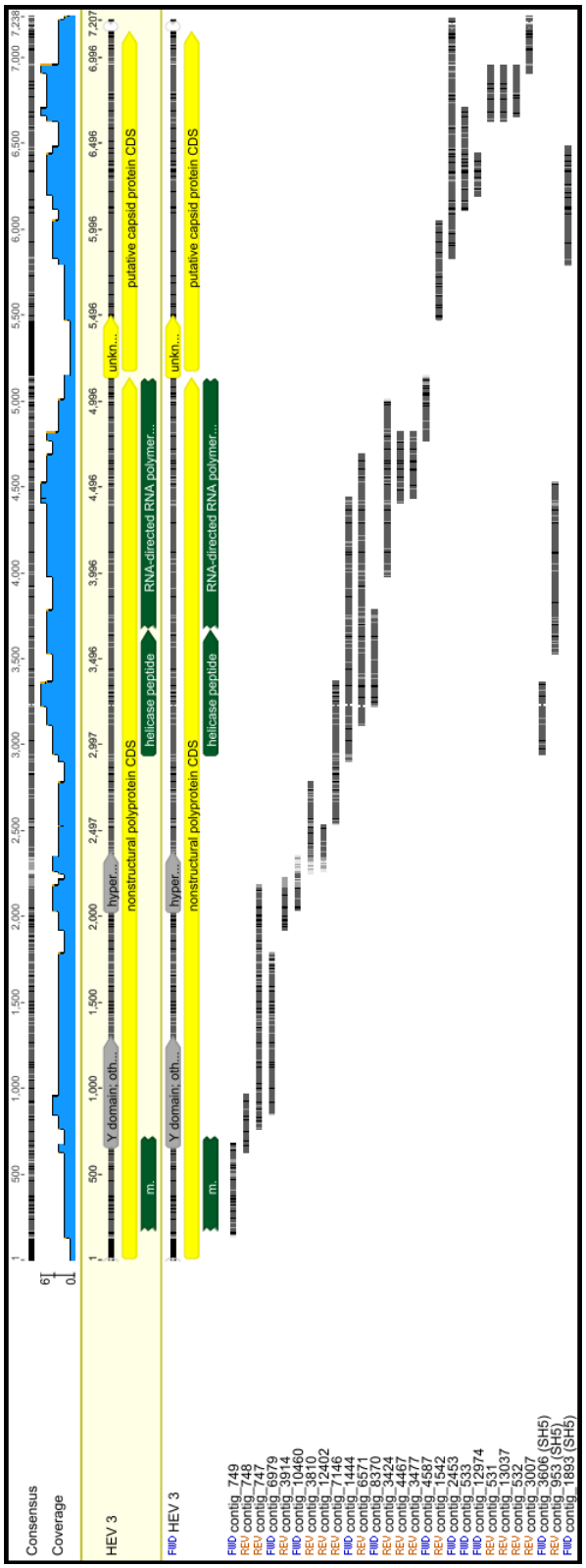


New group Anello

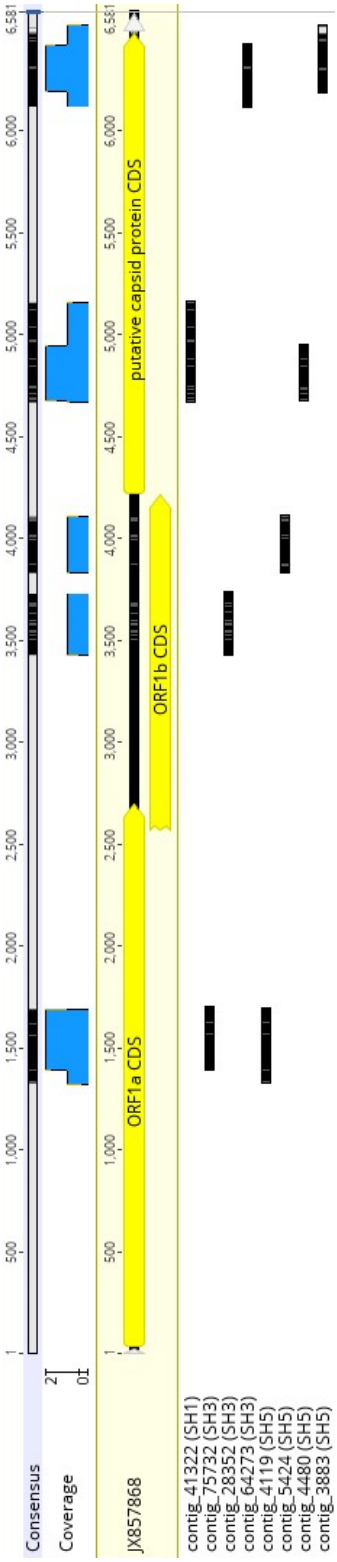
Contig 1199 sh5



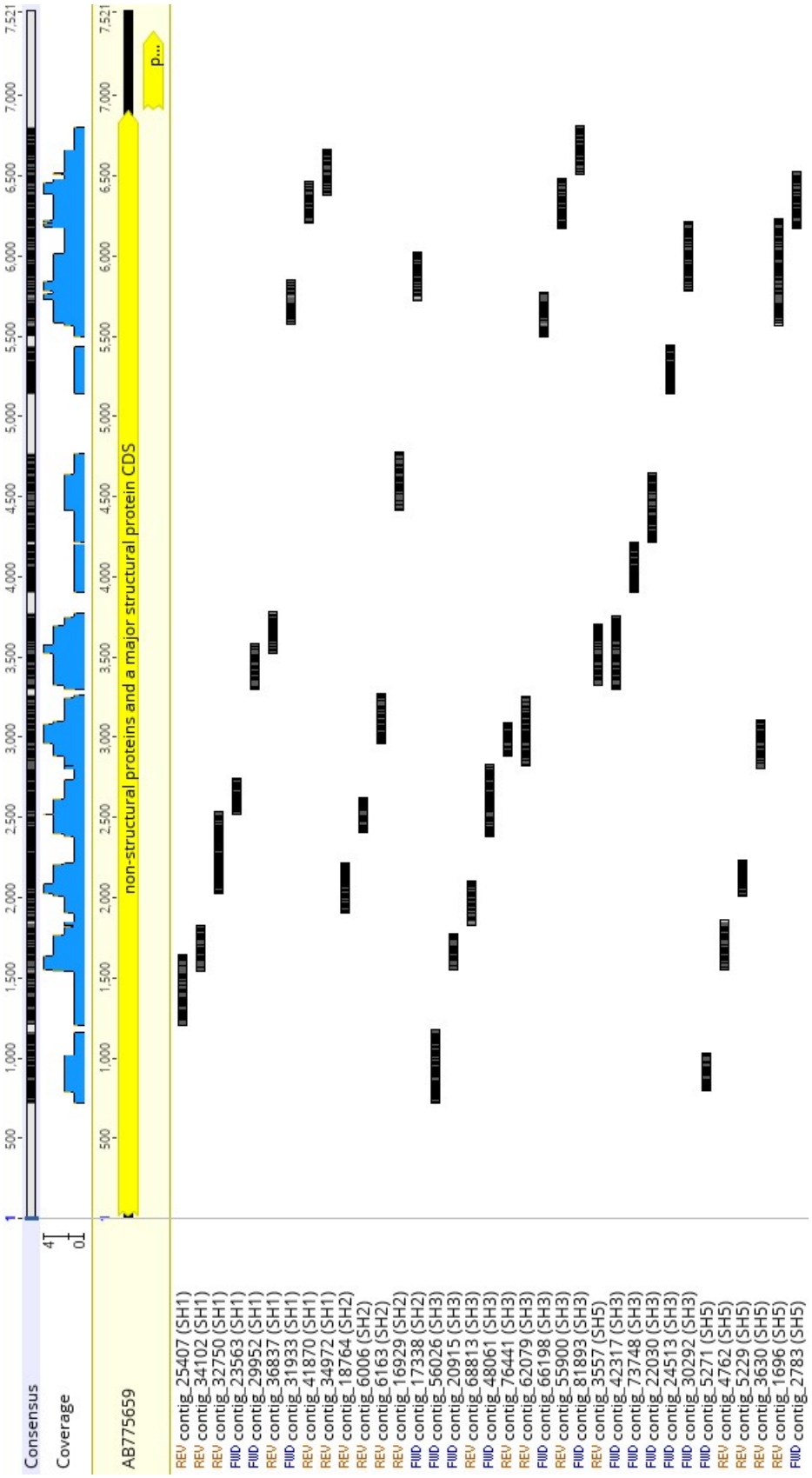
8.3 Alignment of the all the contigs founded in HEV and AI+ImSP pools with the reference complete genome of Hepatitits E genotype 3



8.4 Alignment of the all the contigs founded in Male A, Female and AI+ImSP pools with the reference complete genome of Astrovirus VA3



8.5 Alignment of the all the contigs founded in Male A and B, Female and AI+ImSP pools with the reference complete genome of sapovirus GV.2



9 OTRAS PUBLICACIONES

International Journal of Hygiene and Environmental Health 219 (2016) 405–411

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

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

UV disinfection and flocculation-chlorination sachets to reduce hepatitis E virus in drinking water

Laura Guerrero-Latorre [a], Eloy Gonzales-Gustavson [a], Ayalkibet Hundesa [a], Regina Sommer [b], Girones Rosina [a,*]

[a] Laboratory of virus contaminants of water and food. Department of Microbiology, Faculty of Biology, University of Barcelona, Av. Diagonal 643, 08028 Barcelona, Catalonia, Spain

[b] Center for Pathophysiology, Infectiology and Immunology, Institute for Hygiene and Applied Immunology, Water Hygiene, Medical University Vienna, Kinderspitalgasse 15, 1095 Vienna, Austria

ARTICLE INFO

Article history:
Received 11 February 2016
Received in revised form 5 April 2016
Accepted 5 April 2016

Keywords:
Hepatitis E virus
Water disinfection
UV radiation
Flocculation-chlorination sachets

ABSTRACT

Hepatitis E Virus (HEV) is a major cause of waterborne outbreaks in areas with poor sanitation. As safe water supplies are the keystone for preventing HEV outbreaks, data on the efficacy of disinfection treatments are urgently needed. Here, we evaluated the ability of UV radiation and flocculation-chlorination sachets (FCSs) to reduce HEV in water matrices. The HEV-p6-kernow strain was replicated in the HepG2/C3A cell line, and we quantified genome number using qRT-PCR and infectivity using an immunofluorescence assay (IFA). UV irradiation tests using low-pressure radiation showed inactivation kinetics for HEV of 99.99% with a UV fluence of 232 J/m^2 (IC 95%, 195,02–269,18). Moreover, the FCSs preparations significantly reduced viral concentrations in both water matrices, although the inactivation results were under the baseline of reduction (4.5 LRV) proposed by WHO guidelines.

1. Introduction

Hepatitis E Virus (HEV) is an emerging virus causing water- and food-borne disease of global significance. The World Health Organization (WHO) estimates there are 3 million acute cases of HEV and 56,600 HEV-related deaths per year (WHO, 2014). Although the majority of HEV infections are subclinical, when HEV does cause clinical symptoms, they can have severe consequences, including fulminant hepatic failure and death, most often in pregnant women (Kmush et al., 2015). In addition, extra-hepatic manifestations of HEV have been observed, including neurological injury (Kamar et al., 2011).

According to the International Committee on Taxonomy of Viruses (ICTV), HEV has four classical genotypes (1, 2, 3 and 4) belonging to the *Orthohepevirus* genus, and it includes a diverse array of viral variants that can infect different hosts (primarily mammalian and avian). Genotypes 1 and 2 are strictly human, whereas strains corresponding to genotypes 3 and 4 are zoonotic, with pigs being the primary host (Kamar et al., 2014).

The epidemiology of hepatitis E differs between low- and high-income countries. In areas with poor/limited sanitation and hygiene practices, including large parts of Asia, Africa and South America, HEV has caused medium- to large-sized waterborne outbreaks. Over the last decade, outbreaks have occurred in areas of humanitarian emergencies, such as camps for refugees or internally displaced populations (Boccia et al., 2006; Guerrero-Latorre et al., 2011; Howard et al., 2010). The most recent example was an HEV outbreak that spread across South Sudan between 2012 and 2014, resulting in over 10,000 cases and cross-border infections into neighbouring countries, including 367 cases in South-Sudanese refugee camps in Ethiopia (UNHCR, 2014).

Moreover, the increased prevalence of HEV among populations in high-income countries has been well documented, with sporadic patterns of cases due to zoonotic transmission following consumption of raw meat, close contact with infected animals or hepatic transplantation (Kamar et al., 2012).

The fecal-oral route is the predominant mode of transmission for HEV, and as there are currently no efficient curative therapies for Hepatitis E infection, measures aimed at proper treatment of drinking water, safe disposal of human excreta and improvements to personal hygiene are the keystones for prevention and control of this disease (WHO, 2014).

* Corresponding author.
E-mail address: rgirones@ub.edu (G. Rosina).

http://dx.doi.org/10.1016/j.ijheh.2016.04.002
1438-4639/

International Journal of Food Microbiology 257 (2017) 80–90

Contents lists available at ScienceDirect

International Journal of Food Microbiology

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

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

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

[a] Laboratory of Virus Contaminants of Water and Food, Department of Genetics, Microbiology and Statistics, University of Barcelona, Barcelona, Catalonia, Spain
[b] Computational Genomics Lab, Department of Genetics, Microbiology and Statistics, University of Barcelona, Barcelona, Catalonia, Spain
[c] Institute of Biomedicine (IBUB), University of Barcelona, Barcelona, Catalonia, Spain

ARTICLE INFO

Keywords:
Food metagenomics
Viral pathogens
Food safety
Caliciviridae
Picornaviridae
Next generation sequencing

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 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.

1. Introduction

Food-borne diseases remain 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, food

* Corresponding author.
E-mail address: x.fernandez-cassi@ub.edu (X. Fernandez-Cassi).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.ijfoodmicro.2017.06.001
Received 24 December 2016; Received in revised form 30 May 2017; Accepted 3 June 2017